DOCUMENT RESUME

ED 194 001 HE 013 213

TITLE	Assessment. Current Issues in Higher Education No. 5, 1979.
INSTITUTION	American Association for Higher Education, Washington, D.C.
PUB DATE	79
NOTE	42p.: Papers presented at the Annual Meeting of the American Association for Higher Education (Washington, DC, March 1979).
AVAILABLE FROM	American Association for Higher Education, One Dupont Circle, Suite 780, Washington, DC 20036 ($2.50)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Academic Achievement; *Academic Standards; Accountability; *Accreditation (Institutions); Accrediting Agencies; *Basic Skills; Case Studies; *Educational Assessment; Evaluation Criteria; Higher Education; Institutional Evaluation; Liberal Arts; Minimum Competency Testing; Program Budgeting; Retrenchment; State Programs; *Student College Relationship; *Student Evaluation; Student Experience; Student Needs
IDENTIFIERS	*Program Discontinuance

ABSTRACT

Various topics concerning assessment in higher education are considered in five papers. In "Defining and Measuring the Competencies of a Liberal Arts Education," David G. Winter proposes that liberal arts competencies can be assessed by creative new quantitative strategies and outlines five important criteria, or standards, as well as measures of student evaluation and educational effects. In "Statewide Assessment of Basic Skills," Alice J. Irby considers state government intrusion in higher education and describes the spreading practice of statewide basic skills assessment of students even before they enroll in college. In "Identifying and Evaluating Major Elements in the Quality of Student Life," Thomas R. Plough suggests that once enrolled and in attendance students assess their institution in terms of its maintaining or improving student quality of life. In "What Is the Most Effective Way to Handle Program Discontinuance? Case Studies from 11 Campuses," Edward A. Dougherty claims that program termination may cost more than it saves unless institutions plan carefully and gain support from those affected. Problems and effective strategies in 10 instances of program discontinuance are analyzed. In "How Do You Know a Good College When You Stumble Over One?" John Harris describes how accrediting associations are seeking to broaden their ideas about assessment, emphasizing student outcomes and competence as the bases on which institutions should be held accountable. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Current Issues

In Higher Education

1979

ED194001

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

American Association for Higher Education

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

(5) ASSESSMENT

David G. Winter
Alice J. Irby
Thomas R. Plough
Edward A. Dougherty
John Harris

HE 013 213

American Association for Higher Education

The American Association for Higher Education

The American Association for Higher Education is a national association open to all individuals concerned about the value and effectiveness of higher education. Each year AAHE's National Conference on Higher Education brings together approximately 2,500 educational administrators, faculty, students, government and foundation officials, representatives from accrediting and testing agencies, and others, to clarify issues, disseminate effective educational practices, and discuss the future of higher education.

The theme of the 1979 National Conference was "The Academic Workforce: Opportunities for Renewal." Some of the papers prepared for presentation at the conference were selected by a review committee for publication in the annual conference series *Current Issues in Higher Education*. Many of the papers in this series focus on the quality of worklife in academe and ways to improve it. In some cases, the papers have been revised by their authors and therefore differ somewhat from the actual conference presentation. They have not, however, been extensively edited or revised, since the purpose of the series is to disseminate the conference presentations as rapidly as possible. Points of view and opinions reflected in this series are those of the authors, not the Association.

AAHE Members:

Current Issues in Higher Education, 1979 *will be sent to you weekly during September and October. The seven monographs that you will receive are listed inside the back cover.*

Board of Directors, 1979-1980

President
Martin Kramer, Senior Fellow, Carnegie Council on Policy Studies in Higher Education

Vice President (President-Elect)
Martha E. Church, President, Hood College

Immediate Past President
Wilbert J. McKeachie, Director, Center for Research on Learning and Teaching, The University of Michigan

Members
Alexander W. Astin, Professor of Higher Education, University of California-Los Angeles, and President, Higher Education Research Institute

Marie Childers, Graduate Student, Department of Higher Education, Indiana University

Morton Darrow, Vice President, Planning and Analysis, The Prudential Insurance Company of America

Paul A. Elsner, Chancellor, Maricopa County Community College District

Sister Collette Mahoney, President, Marymount Manhattan College

Arturo Madrid, Associate Dean and Executive Officer, College of Liberal Arts, University of Minnesota

Robert F. Sexton, Executive Director, Office for Experiential Education, University of Kentucky

Donald Stewart, President, Spelman College

Elizabeth Wells, Consultant, Cincinnati, Ohio

Willard Wirtz, Chairman of the Board, National Manpower Institute

Published seven times a year during September and October. Application to mail at second-class postage rates is pending at Washington, DC.

Subscriptions: Subscriptions are $12.50 for the *Current Issues in Higher Education, 1979* series. Nine dollars of the AAHE membership fee covers subscription costs for *Current Issues in Higher Education*. Ordering information and prices for individual monographs are provided inside the back cover.

Copyright 1979, American Association for Higher Education

Assessment

Of the student

Defining and Measuring the Competencies of a Liberal Arts Education, By David G. Winter, p. 1

Statewide Assessment of Basic Skills, by Alice J. Irby, p. 11

By the student

Identifying and Evaluating Major Elements in the Quality of Student Life, by Thomas R. Plough, p. 17

By the institution

What Is the Most Effective Way to Handle Program Discontinuance? Case Studies from 11 Campuses, by Edward A. Dougherty, p. 25

Of the institution

How Do You Know a Good College When You Stumble Over One?, by John Harris, p. 35

Published by the American Association for Higher Education, One Dupont Circle, Suite 780, Washington, DC 20036.

Introduction

Assessment—a problem, a process, a painstaking search for ways to know if our directions are clear, our means appropriate, and our results successful. We examine ourselves and our students and are examined in turn by external groups to whom we are accountable.

How can we do it more effectively, with due regard for accuracy and the full meaning of the ideals we profess?

To what uses will our assessments be put by those to whom we report them?

The papers in this monograph cover a range of assessment topics from student ability to institutional demise.

Some highlights:

- *David Winter's* proposal that liberal arts competencies can be assessed by creative new quantitative strategies—it elicited trenchant opposition at the AAHE conference from liberal arts college president Leon Botstein and sharpened the debate for those present.

- *Alice Irby* raises the spectre of state government usurpation of faculty and institutional prerogatives in her description of the spreading practice of statewide basic skills assessment of students even *before* they enroll in college.

- Once enrolled and in attendance, suggests *Thomas Plough,* students assess their institution in terms of its "efforts to maintain or upgrade student quality of life." A student reactor to his paper pointed out that such quality includes satisfying student expectations about a whole range of student services. The radicals of the 60s have become the determined consumers of the 80s, and the stakes are their very attendance.

- *John Harris* describes how accrediting associations are seeking to broaden their ideas about assessment, emphasizing student outcomes and competence as the grounds on which institutions should be held accountable. Full assessment includes, he says, the judging of an institution against established criteria of academic integrity.

- Program termination, a critical issue in a world of shrinking resources, may cost more than it saves, claims *Edward Dougherty,* unless institutions plan carefully and secure the understanding and cooperation of those affected. His paper analyzes problems and effective strategies in ten instances of program discontinuance.

Taken one by one, these papers provide up-to-date analyses of the complex assessment issues facing institutions of higher education. Taken as a whole, they illustrate a growing awareness in higher education of responsibility to not only effect real change in students, but to do so in settings conducive to learning that operate according to rational processes which do not violate academic ideals. Assessment is higher education's "sensing-device" by which its internal and external workings are examined, its effects known and its purposes justified.

John Bruce Francis
Associate Professor of
Higher Education
SUNY at Buffalo

Defining and Measuring the Competencies of a Liberal Arts Education*

By David G. Winter

For over two thousand years, liberal education has been an educational ideal of the West, for the brightest elite if not for all students. Its roots go back at least to Plato's philosopher, who was trained "to distinguish the idea from the objects which participate in the idea," and who ought to rule because "in no other state can there be happiness private or public."

Since Plato's time, liberal arts education has had its good times and its times of troubles. The British, the Confucians, the Jesuits, and more recently the Americans all ruled far-flung empires with legions of young men trained in a classical curriculum rather than in the technical skills presumably more "relevant" to their actual careers. This classical curriculum—whether as narrow as the grammar of classical Greek and Latin, or as broad as a course in "Western Thought and Institutions"—was thought to instill a sense of proportion, an equilibrium, and a style of confident assertion tempered by self-restraint in the service of group loyalty (see Wilkinson, 1964).

More recently, the liberal arts have come under attack from several different directions. In the 1960s, our liberal arts colleges were blamed for turning the "best and brightest" students into tools of an imperialist establishment. In the more vocationally conscious 1970s, many liberal arts students worry that they are *not* being made into tools of the establishment! The counterculture of the late 1960s argued that society needed to be "deschooled"; that cognitions, words, and "linear" reflection had to give way to emotions, nonverbal expression, and "spontaneous" action. In recent years the liberal arts have also been attacked as elitist, racist, and sexist.

Perhaps the unkindest blows have come from demographic trends and the economy. By now everyone in higher education knows about the approaching dip in the curve of college-age young adults in America. At the same time, persistent cost inflation means both soaring costs and jeopardized material resources and endowments. In higher education, which is very labor intensive (and especially professional-labor intensive), there are limits to cost-cutting and higher "productivity." At what point, then, will the cost of a liberal arts college education become prohibitive to most American families? When will even the most supportive legislators and foundation executives turn to other claims on public and philanthropic funds? Sometimes we wonder whether this great intellectual ideal of the West is about to go the way of the ocean liner and the country estate—comfortable, full of happy memories, elitist, extravagant, expensive to staff—and hopelessly out of date. Perhaps this economic assault, then, is the most dramatic and dangerous one.

If we are ever going to make explicit cost/benefit calculations for liberal arts education, then the "benefits" or presumed effects of liberal education have to be defined in terms that at least approach (or try to approach) the precise, careful objectivity with which costs can be calculated.

The Need for Measures

The questions that are asked in this demographic and economic climate are familiar ones: (1) What effects does liberal arts education have? (2) How do we know that it has these effects? Today, however, we can no longer treat these questions as academic. Nor can these questions be answered by a kind of religious appeal to faith or personal testimony, as Caroline Bird (1975) suggests we usually make. The calls for accountability and accounting in liberal arts education demand a response, and they demand evidence. Beliefs are not enough—whether faculty beliefs that their teaching really has important effects, or student beliefs that they are having a meaningful experience. To date the evidence about the effects of a liberal arts education isn't very impressive. As Feldman and Newcomb (1973) concluded in their review of over thirty years of research on higher education, college—especially the liberal arts—makes students a little more tolerant and interested in intellectual and aesthetic matters, and a little less interested in religion. It is not a strong case for the investment of tens of thousands of dollars from family, endowment, and tax funds per student.

Psychologists, with their often reductionistic style and narrow, simplistic conceptions of how to define and measure complex things, must share a good deal of the blame for this state of affairs. As Victor Butterfield, one of

*Preparation of this paper was supported by Grant OE 300-76-0310 from the Fund for the Improvement of Postsecondary Education and Contract 400-78-0028 from the National Institute of Education, both to McBer and Company.

David G. Winter is Research Director of McBer and Company, Boston, Massachusetts, and Associate Professor of Psychology at Wesleyan University.

the greatest presidents of Wesleyan University, put it over thirty years ago:

> I [spent] a good deal of time attempting to discuss . . . the ends or purposes of the liberal college. . . . Whereas I met many men in all fields who seemed to me to have little interest or insight into what I considered the problem . . . the proportion of such men in the field of psychology . . . is relatively great.

I would like to suggest a fresh start at the task of defining and measuring the competencies of liberal arts education. I have shifted to the word "competencies" deliberately. If we are ever going to make explicit cost/benefit calculations for liberal arts education, then the "benefits" or presumed effects of liberal education have to be defined in terms that at least approach (or try to approach) the precise, careful objectivity with which costs can be calculated.

The competency-based education movement has tried to do this. Traditional "courses" and highly general statements of goals are broken down and then specified in terms of students' behavior or performance. When this articulation is done well, it is argued, assessment and credentialing flow along naturally from definition. The competency perspective has diverse roots: in American pragmatism, in the "scientific management" or efficiency movement in the 1920s, in the theories of John Dewey, and in the wartime demands for quick training of many people in complex skills. (See Grant & Riesman, 1979, for a general discussion of competency-based education.)

For reasons both philosophical and practical, the competency-based approach grates on the sensibilities of many liberal arts professors. They would argue that, yes, we have outcomes; but no, they cannot be "broken down" and measured in the way you would measure the production process for an automobile. Our outcomes are ineffable; and besides, what if the measures show that we aren't *having* any effects? I do not think that this clash of perspectives is necessary, as the experience of Alverno College in pioneering the competency-based perspective in liberal arts education has shown. I do not propose to try to reconcile the two perspectives, but rather I want to explore some of the implications of thinking about the goals (and presumed effects) of liberal arts education as competencies.

Well then, what is a liberal arts education? In terms of curriculum, the exact boundaries are unclear. Different educators would design different courses of study. Is it the study of classical languages, literature, and history—the "Greats" of Oxford, in which the rulers of Empire were schooled? Is it the study of certain "core" disciplines or bodies of knowledge, courses such as "Western Civilization," "Modern Literature," or a particular set of "Great Books"? Is it *any* course that is consciously and deliberately interdisciplinary or multidisciplinary, such as "Freedom and Authority in the Modern Novel," or "Science and Responsibility"? On the other hand, many professors argue that the essence of the liberal arts is not *what* is learned but *how* it is learned—concepts rather than facts, independent inquiry rather than rote learning. Finally, some educators see the liberal arts as a kind of residual category, including everything that is not of obvious practical use or vocational relevance, especially if it is esoteric or trendy: "The Don Juan Theme in Literature and Society," or "Ancient Icelandic Sagas and Modern Rock Ballads."

Although there is some confusion about the definition of a liberal arts education, there is substantial convergence at a general level about two major kinds of outcomes, from Plato down through the recent "core curriculum" report of the Rosovsky Committee at Harvard. (1) Education in the liberal arts is supposed to foster a cluster of *cognitive* abilities such as critical analytic thinking, making conceptual sense out of confusion, recognizing analogies and patterns from the past of our own society and from the present of other societies, and the disciplined self-control to recognize elements of truth wherever they may be found, on all sides of confusing and emotional issues. (2) It is also supposed to foster *personal* qualities such as self-definition and the capacity to lead in a disciplined, mature way. (I think this is what Plato meant by "kings and princes with the spirit and power of philosophy.") These are the two core outcomes. Some institutions would give additional emphasis to qualities such as appreciation of what is beautiful, pursuit of what is good, and perhaps sensitivity and intuition toward one's fellow creatures. But I think there is general agreement on these two main goals and some of the three secondary goals. From the competence perspective, these goals are vague and overlapping, but this is essentially what we are trying to do in a liberal arts college.

Leaving aside the problem of whether change in students should be attributed to our teaching, rather than to their living and working with peers, or simply to maturation, how do we know when a student has changed? That is, how do we measure the competencies of liberal education? What makes a good measure?

Five Standards for Measures of Competence

I want to suggest five important criteria, or standards, which a good measure of liberal arts competencies or outcomes should meet. I will discuss each standard briefly and then go on to look at how well different kinds of measures—both old and new—meet these standards.

(1) A measure of liberal arts competence should seem to measure what it purports to measure; that is, it should have face validity. To many psychologists, face validity is unimportant and even undesirable if it arouses "response set" or faking tendencies ("psyching out" the question or test) in students. To many professors and educators, on the other hand, this requirement is absurdly obvious, for they recognize no other kind of validity. Given the politics of higher education, face validity is necessary for a measure that is going to be taken seriously, even if on psychometric grounds it is not sufficient. Professors simply will not consent to evaluation of their teaching (especially as part of a tenure decision) or their students' learning by measures that do not, in their view, "do justice" to the process. An example: According to Feldman and Newcomb, my classes are supposed to make students more theoretical and perhaps less religious. I-as-research-psychologist might try to assess

this presumed change with "items," for example, this one adapted from one of the most widely used questionnaires in educational research, the Allport-Vernon-Lindzey *Study of Values*:

> Which of the following two headlines would you be more likely to notice and want to read about:
> (Check one)
> (a) Scientist Announces New Theory (theoretical)
> (b) Theologian Discusses New Docrine (religious)

After taking my courses, students should tend to change from checking answer (b) to checking answer (a). But I-as-professor protest: *that* question isn't what I meant at all. Some psychologist, I bitterly complain, has used a meat cleaver where a scalpel was required. Speaking as a professor rather than as a research psychologist, I simply will not consent to have my career and reputation rest on check-mark answers to questions like that. And, along with parents, employers, and maybe even the courts, I take a dim view of that kind of evidence base for my students' grades, degrees, or credentials.

If this example seems valid, it suggests that most traditional psychological tests will not be good measures of liberal arts competencies. They are "objective," quick, easy, cheap, and mindless to score. They do have an established reputation as measures of aptitude or readiness for liberal arts education—a reputation perhaps undeserved and in any case now under severe attack. They may be administered as a part of the normal work of institutional research offices. But as serious measures of the effects of liberal education—measures that could be used to credential students and to evaluate programs—they simply will not work because faculty do not believe in them. Most faculty will insist that measures of effects look like they measure what they claim to measure and that measure outcomes at the same levels of complexity, sophistication and subtlety as those outcomes are taught.

(2) A measure of liberal arts competence should be generic, applicable to different levels and across different fields. This standard conflicts a little with the previous one, because one easy way to enhance the face validity of a measure is to tie it closely to a specific content area. Letting "a thousand measures bloom," however, makes it difficult to establish convergence. For example, is a test of "critical thinking" for experimental social psychology measuring the same thing as the physiological psychology critical thinking test, or the other tests custom-designed for use with macroeconomics, German romantic poetry, or Wittgenstein's philosophy? And how do these measures at College *A* compare with other, different ones at College *B*? The greater the proliferation of outcome measures, then, the greater the work necessary to demonstrate convergence and equivalence (in order to establish fairness). The more generic the measures, the quicker we can move on to satisfy the rest of the standards.

(3) A measure of liberal arts competence should be based on criteria that are public, so that different evaluators would make approximately the same assessments of the same student performance. In psychometric language, this is called "scorer reliability" or "interobserver agreement," but I prefer the term "public criteria" for several reasons. First, two observers can sometimes agree on how to apply what are essentially esoteric, intuitive criteria. Public criteria are necessarily criteria on which observers can agree, but they are much more. They are explicit. They are known in advance. Students, therefore, know the goals toward which they are studying and working, so that liberal learning becomes a little more of a collaboration between student and professor, instead of a hide-and-seek game in which students try to find they-know-not-quite-what that professors hide they-know-not-quite-where. When evaluative criteria are public, then evaluations can be questioned by students and defended by professors in a straightforward fashion that dampens students' suspicions of favoritism, fate or futility.

(4) A measure of liberal arts competence should have educational validity. That is, it should be shown to change as the result of educational programs and interventions. Students at the end of the program should score higher than other students (or than themselves) at the beginning of the program. This standard seems obvious, yet it is often neglected or treated casually in practice. It implies several things. First, it implies careful coordination between curriculum design and measurement of competence. We should try to test that which we have taught; and if we evaluate students' performance with respect to criterion *X*, then we should design our teaching to be sure that we emphasize *X*, either explicitly or implicitly. Second, however, it implies the use of measures that are capable of showing change. Measures that appear to reflect stable "traits"—such as intelligence or abilities (if you think of them as stable), personality (if you think of *that* as stable), or abstract stylistic traits—will therefore not be very useful for measuring the competencies of liberal education. In fact, because the canons of personality and ability testing require that these measures be stable or reliable—that subjects give the same answers on a second testing that they gave on the first testing—they have a built-in bias *against* showing educational change. When measures of this kind are used to evaluate students' performance or program effects, they tend to work against showing any change, not necessarily because there has been no change, but because the component "items" of these measures have been screened so that only the most "stable" (resistant to showing change) survive.

If we accept this fourth standard, then we must realize that many, if not most, traditional measures of ability—such as are used for screening students at entry—are not very useful as measures of learning, competence, or effects.

(5) Finally, a measure of liberal arts competence should be demonstrably relevant to performance in later life. Here the issue is not *whether* the measure is relevant, but the basis for the claim and the kind of evidence brought forward to back the claim. All liberal arts educators believe that their measures of evaluation are "relevant" to later life; the question is whether this belief rests on *a priori* faith, or on some kind of demonstrable evidence about the measures.

Measures of Student Evaluation and Educational Effects

Traditional Examination Questions

Having suggested these five standards for measures of liberal arts competence, I will now look at some of the ways that we actually go about evaluating students, and discuss them in the light of these standards. Let me begin with the most traditional and typical example. At the end of my course in "Personality," I have given a final examination which includes an essay question such as the following: "Compare and contrast the theories of motivation of Freud and Maslow, with special attention to the notion of unconscious motivation." It is late at night, the day after the final grades were already due at the registrar's office, and I am plowing through the exams. I pick up the next bluebook and read it through. My thoughts might run something like this:

> Hmmm. Well, this student says most of the right things about Freud and Maslow. The facts are pretty good. Well, the organization is OK, but it rambles a lot. That whole first-page rambling discourse on Freud being "linear" and "irrelevant" to 20th century America is off the point. Well, all the facts are there, but the organization just isn't up to "A" standards; it's not creative and crisp. But it's better than the usual "B" rehash of what I lectured about in class. I guess I'll give it a "B+." Is it really worth an "A−"? Isn't this as good as that exam I gave an "A−" to an hour ago? Oh my, it's so late and I have so many more exams to do. "B+" it is then.

How does this familiar measure of liberal arts competence stack up against the standards I have just proposed? First, what is the competence being measured? Is it factual knowledge of Freud and Maslow? Apparently so, but there also seems to be some implicit competence involving organization, synthesis, and writing style, at least so far as the threshold between "A" and "B" goes.

Is my exam question face valid? I, as professor, examiner, and grader, certainly believe that it is and I would guess that most of my professorial colleagues would agree. This measure "looks like" it measures students' ability to remember and then use factual knowledge about Freud and Maslow. Certainly it is better than an "objective" question such as: "The unconscious played a more important role in Freud's theory of motivation than it did in Maslow's theory of motivation. True or False?"

Is my exam question generic? It is tied to particular subject matter (Freud and Maslow), but I certainly believe that the ability to marshal, organize, and use facts, which it reflects, is a generic ability, and that a student's answer to this question would be of about the same quality as his or her answer to any other similar essay question, assuming some minimal mastery of the facts in question.

Are the criteria public? Yes and no. I can usually defend my grading to a disappointed student, although I am not very explicit about what I expect, especially in the areas of organization of facts and clarity of writing. Would my colleagues agree with my evaluation? It's an interesting question: normally we don't grade each other's exams, although we do discuss and compare students' performance in different classes. (I remember a sense of elation when a colleague agreed with me that a certain student was "always just on the borderline of B+/A−," and that both of us always opted for B+. Of course, we never bothered to establish explicitly just what this "border between B+ and A−" was!)

Does this measure show change as a result of education—in this case, my personality course? I don't know. I have never tried giving the final exam on the first day of class. Presumably, a student who hadn't learned anything about Freud and Maslow would not write a very good answer, so that the "facts" aspect of the measure does reflect the teaching of my course. On the other hand, I am not at all sure that the "use and organization of facts" or "clarity of writing" aspect of the measure (the A−/B+ borderline area) has anything to do with the personality course, or that it would show change as a result of the course (controlling somehow for factual knowledge). I suspect that within the context of a single course, I tend to treat writing style as a stable trait rather than as something I can change.

The traditional method of measuring liberal arts competencies is a series of private judgments backed by undemonstrated personal confidence about later-life relevance.

What about relevance to later-life performance? Of course I believe, along with most professors, that my assessments of student performance via exam grades (and the letters of recommendation that are based on these and other grades) are full of importance and relevance for students' later lives. How can someone who can't answer my Freud and Maslow question dare to turn out a success in politics, business, or law—let alone as a professor of psychology! Yet there is not very much evidence that college grades—from me or anyone else—have any strong relationship to competence in later life (see Heath, 1977, and McClelland, 1973). More important, I-as-professor am usually not very interested in gathering such evidence systematically. For me, it is enough to remember a few "star" students whom I kept in touch with, who answered my questions well and who went on to be successful.

In summary, the traditional method of measuring liberal arts competencies—typified in my hypothetical but realistic example of exam grading—is *a series of private judgments backed by undemonstrated personal confidence about later-life relevance*. Its strongest point is its face validity; almost by definition, I believe that what I do is an appropriate way of doing what I say I do. With respect to the other criteria, however—especially inter-rater agreement and later-life relevance—the traditional method has problems. This is not to say that my way of grading exams wouldn't or couldn't be replicated by other colleagues, only that I don't know and that I rarely ask the question in a systematic way. This is also not to say that my grades predict nothing worth predicting in later life. Rather, this is to say that the later-life relevance of my

grades rests on my own faith or casual reflections about particular striking students, rather than on any systematic demonstration of such a relationship.

Assessment in a Competence-Based Program

The competency-based movement in higher education stresses the explicit definition of learning goals and the careful breaking down of broad, abstract objectives into specific skills which are in turn measured by students' performance. Often the emphasis on explicit goals is paralleled by an emphasis on innovative and nontraditional ways of measuring outcomes. *Simulation*, for example: students perform in a task-oriented small group discussing a simulated real-life administrative problem; they write letters or give speeches or observe videotaped situations. *Production*, for example: students create a product or a multimedia presentation of a point of view or message. *Performance*: students actually carry out the "real-life" adult behavior for which they have been trained, whether it is making medical or nursing rounds in a hospital, playing a piano concerto, or helping a legal-services client. Students may even take over the *self-assessment* of their own performance. How do these innovative methods stack up against the five standards? (See Grant & Riesman, 1979, *passim* for other examples.)

Face validity is largely a subjective matter. If faculty in a competency-based education program have devised a certain kind of measure, they probably agree that it looks like it measures what it ought to. Because of the nontraditional methods often used, I think that it will be harder and take more time to get agreement (that is, mutual satisfaction with respect to face validity) for competency-based measures.

Are the nontraditional competency-based measures generic? This is a problem. Given the emphasis upon explicitness and measurability, the result is often a series of highly specific criteria for a particular performance. Did the student show critical thinking in a simulated report written on a controversial issue (for example, by stating the main points of each side of an argument in parallel form and stating the underlying issue on which each side differed)? If this report was written in a biology class, is it equivalent to the measure developed by a professor of German literature for articulating the comparison between Goethe and Schiller? And are either of these equivalent to a medical or nursing student's diagnostic comparison of two clinical cases? Obviously some measures are generic and some are not. There is nothing intrinsic about competency-based measures, however, that makes them generic, and the pressures for explicitness and detail may actually work against it. This is not to say that competency-based measures are less generic than the traditional grading that I outlined above (where the generic quality of the method is almost wholly assumed but not demonstrated); only that such a quality cannot be taken for granted.

Are competency-based measures based on public criteria? Yes, if the categories and standards used to make evaluations are explicitly spelled out, broken down, and rigorously defined, especially if there are systematic procedures for training evaluators and monitoring their agreement. On the other hand, there is nothing intrinsically "public" about competency-based measures if the evaluation of a student's performance is a simple, subjective, and undifferentiated "Yes/No" ("Pass/Fail," "Credit/No Credit," or "Meets criterion of competency/Does not meet criterion of competency") made by some internal or external evaluator.

What about validity? Both educational validity and later-life validity or relevance are not *ipso facto* guaranteed just because a measure is competency-based. If measures are systematically given before and after instruction or practice in the relevant competency, then educational validity is demonstrated; but it cannot be assumed just because the competency-based measure, being nontraditional, somehow "looks more relevant" than traditional measures do. Relevance to later life, as well, is not automatic. I may design the most realistic simulation of real-life on-the-job performance that I can; and yet without further evidence, the conclusion that this measure has real-life relevance involves the same leap of faith that I-as-traditional-professor had made about the Freud-versus-Maslow question. No matter how realistic, a simulation is a simulation and not a real-life performance; and college is college and not a career. (Even "real" medical rounds made with an evaluator, or in the knowledge that the evaluator is monitoring the performance, are not the same thing as the "real" rounds of a "real" doctor.) Of course, it seems reasonable to believe that these simulations are "pretty close" to real life, or that they are obviously closer than are traditional methods. (For example, students may say that the new methods "feel" more "relevant.") Such beliefs are, however, just that—beliefs. Of course, evidence can be gathered for any measure, traditional or nontraditional, explicitly competency-based or not, to demonstrate its relevance to real-life careers and roles. My point is that without such evidence, even the measure that "feels" or "seems" most relevant rests on the same faith that the traditional professor has.

My purpose in discussing competency-based measures in this way is not to draw invidious comparisons. Rather, I suggest that there is nothing special about a measure just because it is called "competency-based." The new, fashionable label does not magically solve the problems of evaluating and measuring performance—in the liberal arts or in any more applied context. In philosophical, epistemological, and scientific terms, competency-based measures are not *intrinsically* different from traditional measures, with respect to the five criteria I have presented.

Let me illustrate what I mean. At Alverno College, perhaps the pioneer of competency-based liberal arts education, serious work is proceeding along all five of these fronts (see Alverno College Faculty, 1979; Mentkowski & Doherty, 1979). Face validity, as usual, is almost guaranteed after enough discussion among faculty. Meanwhile, new measures are continually being designed to tap competency attainment at more generic levels. Considerable effort is spent on developing explicit (and lengthy) manuals and continuous checking of agreement among evaluators. Finally, research is underway to explore and document relationships between competency measures and adult performance—both longitudinal studies of students after

college and comparative studies of students and competent performers in the real world. These research and training tasks are time consuming, and in the context of an ongoing institution of higher education and real-life research settings, they cannot always be carried off with the psychometric precision of laboratory experiments. At least in the case of Alverno College, however, there is a commitment to approach each of the five standards.

Some New Measures of Educational Effects

So far, I have tried to suggest that the problems and standards of measuring the competencies of liberal arts education are the same whether the *mode* of evaluation is traditional or nontraditional.

I would like to conclude with a brief discussion of some new measures of liberal arts competencies that were originally developed not to evaluate individual students' performance, but rather to do research on the effects of different kinds of higher education. This research was conducted first by Abigail Stewart of Boston University, David McClelland of Harvard University, and me (Winter, Stewart, & McClelland, 1978); and it is now being conducted by McBer and Company in Boston. We had been concerned about the state of research findings on the effects of higher education, to which I alluded above. Despite decades (or centuries) of luxuriant language about the goals of liberal education, there is little firm evidence about whether these goals are actually achieved. We believed that the evidence to date was more a reflection of the tests used than of the efficacy of liberal arts education. For example, psychologists who study "thinking" or information-processing typically focus on relatively simple elements or processes, things that seem very far away from what a college student does, for example, in answering a question like the following:

> Lord Acton said that all power corrupts, and absolute power corrupts absolutely. Discuss, with reference to Watergate and the Shah of Iran.

We felt that tests of liberal arts competencies should be based on what college students actually *do*, rather than on what psychologists can easily (and cheaply) score. This is, of course, the first standard measured above, that of face validity. If one presumed competence of liberal education is forming and articulating complex concepts, then as a measure of competence students should be given a mass of complex and confusing information and be asked to organize this information into their own coherent concepts, rather than being given examples of somebody else's concepts and asked to check which information applies to which concept—or worse still, examples of some artificial concept devised by a psychologist for laboratory research (e.g., combinations of shapes and colors, or even Chinese characters) and then be asked to recognize further examples of the "concept" and "not concept."

In devising a test of complex concept formation, therefore, we tried to stay as close as possible to real-world tasks. We gave students two groups of brief, imaginative stories and then asked them to formulate and describe the differences between the two groups in any way they liked. We call this the *Test of Thematic Analysis* (see Winter & McClelland, 1978). While the two groups of stories differed in many ways, the idea of the test does not depend on using these stories. Any two groups of complex information—novels, paintings, systems of philosophy, or sets of scientific data—could be substituted for the stories.

We find that most students make rather vague and wandering comparisons, full of their own subjective reactions and emotions. It is as though they are not quite able to master the "blooming, buzzing confusion" of a complex world. Seniors at an excellent, traditional liberal arts college, on the other hand, were able to make crisply worded comparisons, using parallel structure, full of examples and qualifications, and redefining the raw information where appropriate. They were able to recognize things that are functionally similar even though superficially quite different. In short, they are better able to bring the confusion and complexity of daily life under cognitive control. The differences hold up even after adjustment for SAT scores and social class differences among the different groups.

Another complex outcome of liberal education is *intellectual flexibility*—the capacity to see truth in all sides of a complicated and controversial issue. Instead of asking students whether they were intellectually flexible ("always—often—occasionally—seldom—never; check one"), with an obvious invitation to give the "right" answer whatever they actually felt, Abigail Stewart developed a test called *Analysis of Argument*, which actually confronted them with an extreme and badly argued point of view on an emotionally laden issue—for example, an excerpt from a 1968 sermon by Norman Vincent Peale criticizing draft resistance, or from a 1978 speech by a New Hampshire politician blasting student critics of nuclear power plant construction. Students were told to attack the position. Seniors at the liberal arts college made their attack in much more coherent, logical ways, as compared to freshmen at their own college and students at other, more vocationally oriented colleges. (Again, appropriate statistical controls were used to remove any effects of differences in intelligence or social class.) Then, without prior warning, we asked students to *defend* the original argument. The usual reaction was to switch from a blanket condemnation to a blanket endorsement. The liberal arts seniors, however, were able to restrain themselves, to pick out a few elements of truth in the original statement while still being reasonably consistent with the logic of their previous attack.

In these two cases, we began with a major effort to achieve the first standard described above—face validity. I think that these two tests appear to measure what they claim to measure. The detailed manuals for scoring responses to these two new tests are persuasive evidence of their face validity.

Some of our other "new" tests are based on the familiar Thematic Apperception Tests (TAT), and may therefore have less readily apparent face validity. Abigail Stewart devised a measure of *Self-definition* (Stewart & Winter, 1974), which measures freedom from ascribed roles and an instrumental, effective style of thinking and acting, and a measure of *Stage of Adaptation to the Environment* (Stewart, 1977, 1979). David McClelland combined TAT

measures of high power motivation, low affiliation motivation, and high Activity Inhibition (or self-control) into a complex measure of the *Leadership Motivation Pattern* (McClelland, 1975). All of these measures are generic, in the sense that they measure liberal arts competencies across a wide range of different fields or subjects. Contrary to what one might think, for example, the Test of Thematic Analysis does not seem biased in favor of students who are literature or psychology majors, for they do not score higher than do other students.

Each of these measures is based on a detailed coding manual, and our ultimate goal is to have available practice materials so that after brief self-training, scorers can attain 85 percent or better agreement. These new measures, in principle and increasingly in practice, are based on clear-cut public criteria. In fact, these tests could even become the framework for an explicit attempt to teach or shape the appropriate responses to them as a way of "teaching to" the underlying competencies (as McClelland and Winter, 1969, did with the achievement-motivation scoring manuals).

Our initial research study showed that these measures have educational validity. That is, they were designed to reflect changes associated with a liberal arts educational program, and in further cross-validation studies, seniors at a liberal arts college showed scores on each measure that were significantly higher than those of liberal arts freshmen and either seniors or freshmen at other kinds of colleges (a vocationally oriented college and a community college). Research is currently being conducted by McBer (under a grant from the Fund for the Improvement of Postsecondary Education) both to replicate these findings with a longitudinal study of students at the liberal arts college, and also to pinpoint what institutional or eductional features "cause" what changes, through a study of several quite different institutions of higher education.

What about relevance to real-life adult performance, both in career roles and in the management of personal life? There are two sources of data for our new measures. First, several of these measures have established (or growing) construct validity credentials from previous research. For example: McClelland has shown that the leadership motive pattern predicts organizational ability and success among managers, particularly their capacity to generate high morale and performance among subordinates (McClelland, 1975; McClelland & Burnham, 1976.) Self-definition predicts effective handling of stress and life crises (Stewart, 1978). In a recent study of Navy alcoholism counselors (Boyatzis & Burruss, 1978), scores on the Test of Thematic Analysis (modified to consist of two groups of case studies rather than two groups of imaginative stories) predicted success in dealing with clients—i.e., the proportion of alcoholic clients who improved. It seems, therefore, that the cognitive competency of complex concept formation plays an important role even in what might be thought of as an interpersonal role that requires empathy. Apparently you have to create the right constructs in order to help people effectively.

The second source of real-life data comes from a longitudinal study of class-of-1964 graduates from the traditional liberal arts college mentioned above. They had taken the TAT as freshmen (in September 1960) and were followed up as adults ten years after college education, in 1974. (Of course they didn't take the Test of Thematic Analysis or the Analysis of Argument, because these hadn't been developed in 1960-64.) Over this 14-year span of life, the liberal arts competency measures based on even the freshmen-year TAT showed important relationships to real-life adult performance and life satisfaction outcomes, holding constant other relevant background factors of course. For example: the Leadership-Motive Pattern predicted being active as an officer in organizations; the Stage-of-Adaptation measure predicted career satisfaction, being an entrepreneur, and organizational membership—signs of mature adaptation to adulthood; and Self-Definition predicted feeling energetic, having more income, and having a dual-career family.

The results are especially interesting if we compare the predictive power of these new TAT-based measures to that of two traditional measures of "competency," SAT scores and college grades (level of honors at graduation). Both of the traditional measures were actually negatively correlated with many aspects of later career success and satisfaction, confirming at this college the results that Heath found at another traditional liberal arts college and McClelland's general review of the poor power of SAT scores (and college grades) to predict to important real-life outcomes. In terms of the positive predictions, summing across all the important life outcomes we could measure, the new TAT-based liberal arts competency measures accounted for almost 50 percent more variation than did SAT scores and level of honors at graduation combined. Overall, the 14-year longitudinal correlations (or regression coefficients) were significant, although not very high for any of the measures. This is natural, since these crude measures of life outcomes within such a heterogeneous group of young adults are affected by a variety of situational and historical factors, as well as by personality traits and educational competencies other than what we have measured. In this study, however, the real-life-relevance credentials of the new competency measures are substantially more impressive than are those of the two old standbys, SAT scores and college academic performance.

Figure 1 brings together the discussion of the three examples I have given, in terms of what are perhaps the two most important standards: *public criteria* (#3) and *demonstrated real-life evidence* (#5). In the lower left-hand corner, we see my typical professorial reading of an essay as a private judgment whose relevance to the student's later life performance rests on my undemonstrated faith. In this cell also are those competency-based assessments and other nontraditional measures which call for global judgments by observers and assessors, without providing any detailed criteria or training, and which have only a presumed relevance (even if it is an "obvious" relevance) to later life.

The upper-left cell contains measures that involve public criteria, but which have not yet attained demonstrable life relevance. Many traditional outcome

measures, such as SAT and GRE scores, also belong here. In the upper-right cell are those measures that have met both standards. Movement toward the upper-right quadrant, then, is desirable in terms of the point of view I have advanced here, because this is movement toward attaining the two most critical standards for evaluating any measure of liberal arts competencies and outcomes. The work at Alverno College to which I alluded earlier is movement in this direction. To the extent that the assessment criteria are reasonably "public" and trainable (which is probably the easier standard of the two to meet), then the current effort is to move from the upper-left to the upper-right quadrant. Some of the new tests I have described are perhaps already there, while others are based on public criteria but await further demonstration of real-life relevance—i.e., rightward motion in terms of Figure 1.

What is in the lower-right cell? Can any measure have demonstrated validity or relevance to real life without being based on public criteria? I think so. There are probably some professors who, in their intuitive (and private) wisdom, can pick students who will turn out to be the scholars, business achievers, or political leaders of tomorrow. Their letters of recommendation are eagerly sought after by students and are taken seriously by graduate schools, scholarship committees, and corporate executives. If we only knew the criteria they used for making their judgments! Here also would go the interview situations which have a demonstrable capacity to pick those who will succeed, but which are not based on any clearly articulated cues or criteria. Here also would go intuitively made clinical judgments for which there is, nevertheless, evidence of relevance or, in other words, validity (see McClelland & Winter, 1969, p. 268).

Of course, we hope that the new measures I described above, and others like them, can be developed more fully as measures of both individual student performance and of the aggregate effects of educational institutions. I see no reason, for example, why a *Barron's Profiles of American Colleges* or a *Lovejoy's College Guide* of the future should not show comparative profiles of college *change effects* on a standard series of competency measures that are face-valid, generic, reliably scored according to public criteria, and relevant to both educational and "world-of-work" goals. Perhaps federal and state governments, more and more concerned with the careful stewardship of tax money in an era of competing claims for resources, should require that each recipient college and university produce, maintain, and follow an "affirmative documentation plan" to demonstrate their effects, as a minimum standard of institutional accountability, not to say of truth in advertising.

At a more modest level, we hope that the dissemination of these measures, which derive from psychological research, may converge with the efforts of both traditional and nontraditional liberal arts educators

Figure 1. Classification of Several Methods of Measuring Liberal Arts Competencies and Outcomes Along Two Dimensions or Standards

Criteria Used in Making Evaluation	Relevance to Later Life: Assumed; based on personal faith or persuasion	Relevance to Later Life: Based on demonstrated evidence
Public and Trainable	Competency measures with detailed, clear-cut criteria and intensive assessor training SAT and GRE scores Some new measures (e.g., Analysis of Argument), where there is no evidence yet about real-life relevance	Competency measures with public criteria *and* empirically demonstrated relationship to real-life outcomes Some new measures (e.g., Self-definition, Stage of Adaptation, perhaps Thematic Analysis)
Private or Intuitive	Traditional professor's grading of papers and exams Many competency assessment measures	Letters of recommendation from a professor with an established track record as a good "talent scout" Clinical judgments with established validity credentials

to improve the standards used to assess the competencies of education in the liberal arts.

We believe in the liberal education ideal. We believe that our research studies make at least a *prima facie* case that it has some of the effects it claims to have, at least at one institution and probably at many others. But the whole spirit of my argument is that we should not use this one study and this one *prima facie* case as an excuse to lapse back into mere faith or personal testimony as the foundation for our belief in the liberal arts. A better response, with better evidence, will be needed during the demographic and economic time of troubles in the next decades. If we are to defend and maintain the liberal arts ideal, should we ever try to do it with less than our best response and our best evidence?

References

Alverno College Faculty. *Assessment at Alverno College.* Milwaukee, WI: Alverno Productions, 1979.

Bird, C. *The Case Against College.* New York: David McKay, 1975.

Boyatzis, R. E., & Burruss, J. A. *Validation of a Competency Model for Alcoholism Counselors.* A final report submitted to the U.S. Naval Alcohol Rehabilitation Center, San Diego, CA. Boston: McBer and Company, 1978.

Feldman, K. A., & Newcomb, T. M. *The Impact of College on Students.* San Francisco: Jossey-Bass, 1973.

Grant, G., & Riesman, D. (Eds.) *On Competence.* San Francisco: Jossey-Bass, 1979.

Heath, D. H. "Academic Predictors of Adult Maturity and Competencies." *Journal of Higher Education,* 1977, *48,* 613-32.

McClelland, D. C. "Testing for Competence Rather Than for 'Intelligence.'" *American Psychologist,* January 1973, 28.

McClelland, D. C. *Power: The Inner Experience.* New York: Irvington, Halstead-Wiley, 1975.

McClelland, D. C., & Burnham, D. H. "Power is the Great Motivator." *Harvard Business Review,* 1976, *54* (2), 100-110.

McClelland, D. C., & Winter, D. G. *Motivating Economic Achievement.* New York: The Free Press, 1969.

Mentkowski, M. L., & Doherty, A. *Careering After College: Establishing the Validity of Abilities Learned in College for Later Success.* Progress Report on Grant NIE-G-77-0058 submitted to the National institute of Education. Milwaukee, WI: Alverno College, January 1979.

Stewart, A. J. *Scoring Manual for Stages of Psychological Adaptation to the Environment.* Unpublished paper, Boston University, 1977.

Stewart, A. J. "A Longitudinal Study of Coping Styles in Self-Defining and Socially Defined Women." *Journal of Consulting and Clinical Psychology,* 1978, *46,* 1079-1084.

Stewart, A. J. *Psychological Adaptation to Life Change in Children and Adults.* Symposium at the Eastern Psychological Association Convention, 1979.

Stewart, A. J., & Winter, D. G. "Self-Definition and Social Definition in Women." *Journal of Personality,* 1974, *42,* 238-259.

Wilkinson, R. *Gentlemanly Power: British Leadership and the Public School Tradition.* London: Oxford University Press, 1964.

Winter, D. G., & McClelland, D. C. "Thematic Analysis: An Empirically Derived Measure of the Effects of Liberal Arts Education." *Journal of Educational Psychology,* 1978, *70,* 8-16.

Winter, D. G., Stewart, A. J., & McClelland, D. C. "Grading the Effects of a Liberal Arts Education." *Psychology Today,* 1978, *12*(4), 68-74, 106.

Statewide Assessment Of Basic Skills

By Alice J. Irby

Background

Summer 1977
After two years of soul-searching, conversation, and research, the Panel[1] chaired by Willard Wirtz on the SAT score decline concludes that the decline, in part at least, was attributable to pervasive factors in American society involving the schools, the family, the media, and the community.

Spring 1978
The New York Times reports that law firms are concerned about the lack of basic writing skills on the part of the recent graduates of the nation's law schools.

Spring 1978
The publicity surrounding Harvard's "reform" of the undergraduate curriculum focused attention on the core curriculum which, to many, meant "basic education."

Spring 1979
The University of Wisconsin–Parkside is implementing a program of proficiency examinations in reading, writing, mathematics, library skills, and research-paper writing to insure competence in these areas before junior-level work is undertaken. Students who do not measure up are put on probation.

Spring 1979
Mr. Califano, Mr. Boyer and Ms. Graham testify during Senator Eagleton's Subcommittee on Education hearing on basic skills. Ms. Graham cites the National Assessment of Educational Progress reports documenting declines in the quality of essays written by 17-year-olds, and notes NIE's commitment to support projects which may lead to more productive learning and teaching in the area of writing.

Spring 1979
The Education Commission of the States reports that legislation is pending or has been passed in Alabama, Arizona, Arkansas, Colorado, and Missouri which requires competency tests in English and/or mathematics for teachers.

This quick look back a couple of years suggests, along with numerous other articles and statements, that we may be in the middle of a "Back to Basics" movement in higher education which parallels the minimum competency push at the elementary/secondary level. Notwithstanding the fact that most thrusts described as movements in education turn out to be short-lived fads, a number of states and many institutions are struggling (sometimes floundering) to cope with basic-skills issues—competency in reading, writing and computing sufficient to benefit from and prosper in liberal arts and professional programs in our colleges and universities. These efforts take many forms: resolutions of state boards of higher education, locally initiated remedial programs, system-wide or statewide testing of entering students, system-wide or institutionally based proficiency testing at the end of the freshman or sophomore year, and legislated statewide "competency testing" of teachers.

This paper will focus on system or statewide testing at entry into higher education, some of the issues surrounding it, and the experience of one state.

Easy Scapegoats, Easy Answers

When there is a pervasive sense of failure among a population, it often happens that people look for, invent, or take comfort in finding scapegoats. Like the furor that gripped the nation after *sputnik*, which had its fallout on the colleges, the anguish over deficiencies in basic skills among high school graduates is laid at the doorsteps of the nation's schools. Reading, writing and arithmetic are minimum responsibilities of the schools, so the argument goes. We expect the schools to be more than the reflection of the values of the society. So, when trouble comes, there is an institution outside ourselves and our jurisdiction to absorb the losses and institute reform. That lowered parental expectations, reduced motivation among young learners, disinterest of community groups, and emptiness of a TV-viewing, spectator public might dampen curiosity, decrease initiative, and contribute to intellectual shabbiness is ignored or forgotten. So, first the teachers are blamed, and then the teachers who teach the teachers, because the teachers are not prepared to teach the three R's.

The answers many recommend are as easy to find as the scapegoats. I will describe a few here. In 1976, a Gallup poll revealed that 65 percent of those polled felt that high school students should be required to pass a standardized test to graduate. Now there are 36 states which have enacted some kind of minimum-competency-testing legislation, 17 of which require passing tests to receive the high school diploma.

Because the teachers and the teachers' teachers have not maintained standards, so it is claimed, and because subjective standards are variable from place to place and

[1] *On Further Examination*, Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline, College Entrance Examination Board, New York, 1977, pp. 1, 25-37.

Alice J. Irby is Vice President of the Educational Testing Service.

time to time, the search for a strategy to remedy the decline, to reverse the failure, takes the form of looking for absolutes—something that we can grab onto, get a handle on. The answer? Local tests, statewide tests, even national tests. National figures call for national testing programs. Observe the speeches and writings of Senator Pell, Senator McGovern, Admiral Rickover.

The proposed solutions take other forms. Some believe sincerely that a) there are too many unmotivated and ill-prepared students in college and b) that universities should not (and cannot) provide adequate remedial assistance.

Solution #1—Students who cannot handle traditional freshman courses in English, foreign language, science and mathematics should be denied admission. How? By raising admissions standards, specifically by setting minimum scores on the SAT, ACT or similar exam. Sometimes the leaders who call for raising admissions standards are the same leaders who encouraged, pressured and admonished universities to open their doors in the late 60s and early 70s to the nontraditional populations, especially minorities from urban schools.

Solution #2—Stop funding remedial programs at the college level. It's too late in a student's career; the courses are likely to be ineffective. Moreover, college faculties are neither trained nor motivated to provide remedial instruction. Their rewards (and their promotions!) lie elsewhere. Besides, it's much too expensive, both because the extent of remediation required is enormous and the salaries of university professors result in abnormally high unit costs. Finally, college credit should be given only for college-level work. To appropriate scarce resources for any but college-level pedagogy and research results in a weakening of the central thrust and mission of the university.

So, through one or a combination of these approaches —minimum competency testing at the elementary/secondary level, teating at entry to affect enrollments, shifting resources away from remedial support services—the problem is placed back where it belongs (i.e., on the schools), standards in higher education are enhanced, and resources are conserved.

A Closer Look at the Issues and the Players

Would that major societal traumas begot such straightforward solutions with no adverse side effects or repercussions for other goals of a pluralistic society. But fundamental educational change is not easy and complex problems defy simple solutions.

First, a look at the decisionmakers. Whereas most of the initiative for minimum-competency testing at the elementary/secondary level came from outside the educational community and took form through legislation at the state level, such has not been the case in higher education. Exception: proficiency testing for entry into teacher education and for teacher certification. The impetus for testing at the college and university level came from one of two sources, depending upon the state. In Georgia and in New Jersey, the requirement was placed on students and institutions by a state-level policy board. In Georgia, it was the governing Board of Regents of the University System. In New Jersey, it was the coordinating Board of Higher Education, the Chancellor of which is a Cabinet officer. In California, the testing includes the California State University and Colleges but not the University of California. The requirement stemmed from concern among English faculty over the deficiencies of entering freshman students, and the policy board of the system authorized the requirement.

In other states, some universities have undertaken the development of testing programs in the absence of mandates from the system-wide boards of statewide coordinating boards. Placement tests are nothing new and, historically, have been administered by academic departments, especially English and mathematics, prior to matriculation. What is different from the practice of recent years is the focus on skills which students should have prior to matriculation, e.g., basic skills however defined, and the imposition of university-wide or system-wide tests to determine the level of basic skills among the population, either at entry or somewhere along the way. Examples of university-wide programs: Wisconsin–Parkside's certification program mentioned earlier, and the pending recommendation of the Academic Senate Committee on Reading and Comprehension of the University of California–Berkeley that a standardized test be given after completion of required reading and composition courses, regardless of the grade a student receives. Students who fail would have to repeat the course, even if they passed it the first time.

Massive testing of entering students may result in unnecessary testing. It may duplicate internal testing programs.

Historically, faculties of individual universities, through delegated power granted by internal governing bodies, have exercised authority over the academic dimensions of educational programs, to wit: a) the admission of students, b) the design of curriculum, c) the requirements for the degree and the grading standards by which the degree will be conferred, and d) the recruitment of faculty. For some public universities, general admissions policies have been set external to individual units by multi-unit governing boards; in some states, the legislature has played a role from time to time. For the most part, however, the question of what will be taught and who will be placed where has been left to collegiate units, within an existing framework of institutional mission. To the extent that the basic-skills testing programs invade these traditional jurisdictions and impose decisions made external to the faculties of institutions, the risks of short-term dissension and failure increase.

What are some of these risks?

- Massive testing of entering students may result in unnecessary testing. It may duplicate internal testing programs.
- The tests may be inappropriate for many students,

depending upon the heterogeneity of the student population and the diversity of the curricula to which the tests are, ostensibly, related.

- Faculties may lack adequate empirical or clinical data with which to make good use of the test scores and results, in which case the testing serves no useful educational purpose.
- Eagerness to mount programs quickly can result in the production of test results without adequate placement criteria against which they can be used wisely.
- There may be attempts to use a single instrument for multiple purposes, the purposes not being clearly delineated. This, in turn, may lead to inappropriate measurement, in part, and/or the misuse of results, especially in the reporting of aggregate data.
- The effort may receive reluctant compliance at best, or it may be ignored.
- Finally, the entire effort is likely to be costly, with the incremental gain from such a large-scale effort unknown. Of course, even though the educational gains may be unknown or limited, there may be other purposes served which must be recognized in assessing investment and alternate deployment of resources.

Countervailing Views: An Example

New Jersey is the only state which tests all entering freshmen in all public institutions of higher education (the junior colleges, the state colleges, and all units of the university) with a single test in English and mathematics. The 1979 information bulletin calls it a *placement* test. It lasts 3 hours and 15 minutes.

The advocates of the program, who have labored tirelessly to implement the testing program, see a number of advantages:

(1) *Students*—The tests provide information to students about their strengths and weaknesses so they can plan their education program.

(2) *Colleges*—a) The tests help place students in appropriate entry-level courses. b) They also help colleges design and evaluate curricula developed to meet student needs.

(3) *State Education Officials*—The tests provide information on skills deficiencies so that assistance, both technical and financial, can be made available to strengthen remedial efforts.

(4) *Secondary School Officials*—The tests provide aggregate data for understanding the skills levels of their college-bound students and serve as a basis for improved articulation between schools and colleges.

Noteworthy Developments

Since the implementation of the program, several noteworthy items have appeared in the press or in reports issued by the Basic Skills Council.

- Using a *cutoff* scaled score of 65 (on a 35-95 scale with a mean of 65), the Council reported aggregate data indicating that 42 percent of the students tested fell below the 65 on the total English score. The report reads "Appendix D summarizes the need for remediation and assumes a cutoff of 65 on the Reading Comprehension, Sentence Structure, Logical Relationships, Composition, Total English, Computation, and Elementary Algebra sections of the NJCBSPT."

"Students whose general test scores (Composition and Total English) are above a scaled score of 65 probably possess an adequate grasp of those basic verbal skills necessary for success in regular first-year college-level courses. Those students who achieve general test scores below a scaled score of 65 may be deficient in one or more areas of verbal skills."

- The Basic Skills Council is concerned that many part-time students did not take the test. In order to cast a broader net, clarification of the candidate pool has been given to institutions: all full- and part-time freshmen; any freshman transfer student who has not been tested regardless of number of courses completed satisfactorily); and nondegree students who register for a course which would result in accumulation of 12 or more credits (regardless of how well the student performed on the first courses). However, one group is exempt—those first-time students in an institution who already hold an undergraduate degree.
- The testing program is having its impact on resource allocation through the budget process. The basic-skills funding ratio for fiscal year 1980 is 15 to 1 (compared with 21 to 1 for regular lower-division courses) provided: a) the courses are designated as remedial in official college publications, b) the courses do not carry graduation credit and cover noncollege work, and c) placement is made on the basis of clearly stated criteria, which must include at a minimum results of the basic-skills test.
- *The Newark Star Ledger* (Jan. 22, 1979) reported that New Jersey officials expect to eliminate state college remedial programs for students deficient in reading and math skills beginning in 1983 or 1984.
- The Council reported aggregate data by high school for any school with 25 or more students taking the test and expressed its concern about how such data might be misused. Nonetheless, it believes the data can be used constructively and urges the schools to do so, though it does not specify how.
- Finally, the director of the state program is quoted in a recent College Board publication as saying that by testing 43,000 freshmen, New Jersey colleges were able to place more than 99 percent of the freshmen in appropriate language and math courses.

While there may be some students who might learn from the test, there are thousands for whom it is inappropriate, unnecessary, and even insulting.

To many, these developments seem promising: educational reform to insure the success of open access; the adequate funding of remedial courses; the development of basic skills so that students may learn other modes of

communication and open new doors of understanding; the standardization of what is "basic"; and effective pressure on the schools to shape up.

Opposition to the Program

But there's another side to this story that is not as likely to be spread on the public record, because the arguments are subtle and the concerns not easily converted into popular appeals. Those who view the statewide program differently make the following points:

- While there may be some students who might learn from the test, there are thousands for whom it is inappropriate, unnecessary, and even insulting. Why? Because the test is designed to discriminate at the lower end of the SAT (or ACT) scale, not equally or adequately across the scale. Yet students with excellent records at admission, high test scores, and good grades are required to take it. Some institutions petitioned for an exemption policy and understood one was promised subsequently. There is a feeling of betrayal among some now that the exemption score has been set at SAT-V of 600 and SAT-M of 650—which, of course, means that there's an exemption policy in name only.

- "In name only" is also the characterization given to the title of the test itself: *placement test.* Some who are skilled in measurement ask, How can one three-hour placement test in basic skills be effective in placing *all* freshmen in *all* public institutions, especially considering how diverse English and mathematics courses are across each state?

Moreover, how does one flunk a placement test? Yet, some supporters of the program reported that 42 percent of the freshmen "flunked" the test—i.e., scored below 65. The scale (35-95), though designed to be different from the SAT scale, was apparently so similar to many high school numerical scales (0-100) in the public's mind that confusion abounded concerning the meaning of the 65 below which students "failed." The report of the Council itself would contribute to such confusion and misunderstanding. The misuse of data had been a major initial concern of the critics of the program.

- If the test is to serve the placement needs of students and faculty, with institutionally derived standards for use, what purpose is served by the Council's describing results in terms of a statewide cutoff of 65? This contributed to a further sense of betrayal on the part of some who had argued for a state-sponsored test for voluntary use by institutions, for they had operated on an understanding that statewide cutoffs would not be used. Further, they were dismayed when the Council's report to the Board of Higher Education contained such assertions, apparently with no evidence or justification for using a score of 65 to equate with proficiency. Why 65? Why not 60, 50, 70? Simply because the mean of the scale was set at 65?

- Not all faculty and deans found the test helpful. They complained that it was too long, too elementary for their students, unrelated to their curriculums, and unused because it duplicated or disrupted their successful procedures of the past. It gave them no additional information on which to improve their placements. They already knew, as a result of the admissions process, who would need further testing and further assistance.

Some colleges, especially the community colleges, felt forced to abandon effective testing programs. One such program was regarded nationally as exemplary.

Spokesmen for some institutions made the following criticisms in formal protests:

a) Administration and scoring processes do not fit the needs of institutions with rolling admissions and placement programs.

b) The test can be shortened with no loss of measurement.

c) Evidence to support the Council's statements is lacking. The Council refers to levels of proficiency expected; yet no efforts were undertaken to determine standards necessary for academic success.

d) The Council has made no effort to correct the use of the term "failure" by the media and by advocates of the test. Simple comparisons of mean scores are misleading.

e) Just because part-time students have been unable to take the test does not mean they are not receiving equitable services.

- Numerous public school officials not only found the data they received useless, but sharply criticized the Council for reporting misleading data and contributing to public misunderstanding. One must remember that after testing New Jersey freshmen in public institutions, fewer than half the graduating seniors are tested. Additionally, 40 percent of the college-bound students go out of state, many of whom are among the most able. Thus, in no way can the data aggregated by school be representative of school populations. Bias was inevitable; bias was the meat of the media messages.

Some Final Speculations

The conclusion I draw from this example is that the educational benefits of the program are not yet obvious, the political impact may or may not be constructive, and only time will tell whether the program will be undermined, improved to serve institutions and students, or wither away.

If one assumes that the factors at play in this example have surfaced or are likely to surface in other states which undertake mandated statewide programs, one can catalog some claims and fears and estimate some probable outcomes.

At best an externally mandated statewide program can:

- Focus attention on significant educational problems, both among faculty groups and policy makers in positions of power and influence.
- Provide an empirical basis for short-term financial support for remediation.
- Prompt the secondary schools to remedy deficiencies prior to graduation.
- Keep the issues before the public.
- Provide a test or battery of tests for some students and some institutions where none exists or where outside assistance can be beneficial.

Recommendations

John R. Hills provides a good discussion of critical

issues and possible harmful consequences of proficiency testing in a 1977 publication of the Southern Regional Education Board.[2] I recommend it and add a few of my own suggestions below.

(1) Sudden converts to the apparent simplicity of standardized testing can stumble into the trap of believing that because some testing is good, more is better. Following initial testing, a next step is to advocate post-testing to certify either the achievement of basic skills, the effectiveness of remedial programs, or the adequacy of placement services. Though these functions are quite different, the naive true believer may conclude that all these good things can be done with merely a second form of the same test.

(2) Remedial work is not new in public higher education, even among the land-grant universities. Yet, one underlying political purpose of basic-skills testing programs at entry to higher education may be to push all remedial work out of postsecondary institutions—or to relegate remedial courses to some institutions, eliminating it in others. To the extent that funding formulae are used to change institutional functions and to redistribute such resources among institutions, the traditional faculty and institutional prerogatives of determining academic programs and admissions policy are undermined or compromised. In fact, the combination of empirical data derived from basic-skills testing, coupled with budgetary authority, can be used indirectly to set admissions policies statewide among institutions, or as a means of managing retrenchment. This, in turn, can create dislocations among faculties and tensions among institutions during a time of general enrollment decline and belt tightening.

Among educational leaders, there has been consensus that political interference in the academy serves no positive educational purpose. In fact, its long-term effects are likely to be harmful. Whether the repercussions from statewide testing programs extend to the core of the academy depends, in large part, on the educational leadership in the state.

(3) Should remedial courses be banished from the college curriculum, there is the possibility that student populations would become more homogeneous, not just in their basic-skills proficiency. An institution might be faced with racial "resegregation" of its student body, after having toiled to enroll minority students. Many, though by no means all minority students have needed additional work in basic skills. This change is not likely to occur through overt policies; rather, it will occur through changes in admissions practices to enroll only those students who have "adequate basic skills" or through a revolving door of failing those admitted who cannot keep up with the requirements of college-level work the first semester or year. Such trends may result in reintroducing into the institutions of higher education the social stratification that many have fought to overcome.

(4) Finally, undue emphasis on testing can lead to distortions in educational curriculums. Teaching may be oriented toward good performance on the test rather than on a full range of competencies or on developing remedial syllabi appropriate for particular students. Or, if students do well in courses but not on a single test, the difficulty may lie in the test, not in the faculty. Or, if arbitrary judgments and decisions are made with a single test score, absent research, and clinical experience, neither the placement nor the remediation will serve anyone well.

In summary, my own guess is that in the next few years few states will mount statewide basic-skills testing programs at entry to higher education. If certification of high school graduates is the concern, the statewide testing program will normally be administered in the secondary schools. If placement is the issue, institutional variability will likely be the pattern. Certification of basic skills at the collegiate level is likely to be focused at the institutional level (single or multi-campus) at the end of the freshman or sophomore year, even though some state agencies or boards may pass resolutions urging or requiring that something be done. Of course, the scramble for more students may lead to more institutional testing at entry and along the way; or, decreases in funding may lead to faculties' opting to abolish remedial programs (and the instructors who teach them) and less testing. Good testing programs, either for placement or certification, are expensive; even bad ones may be too expensive during the next decade. If reason prevails, we will realize that while state agencies can champion important political causes, effective placement is an institutional matter. Some faculties may not do it well, but removing it from their jurisdiction is not likely to make matters better. An ill-conceived statewide program might nail down a few elusive things with its sledgehammer, but it also tends to flatten everything at the same time.

[2]John R. Hills, *Proficiency Testing: Implication for Higher Education*, Southern Regional Education Board, Atlanta, Georgia, 1977.

References

"All Public Colleges in N.J. Use College Board Services to Assess Freshman Skills," *The College Board News*, February, 1979.

Braun, Robert J., "Grads of 'Best' High Schools Show Need for Remedial Aid," *Newark Star Ledger*, December 10, 1978.

Guskin, Alan E. and Greensbaum, Ben, "Quality and Equality: Skill Proficiency Requirements in the University," paper delivered at American Association of State Colleges and Universities Annual Meeting, November 20, 1978.

Hills, John R., *Proficiency Testing: Implications for Higher Education*, Southern Regional Education Board, Atlanta, Georgia, 1977.

Information Bulletin: 1979, New Jersey College Basic Skills Placement Test, Department of Higher Education, 1979.

Legislative Review, Education Commission of the States, February 19, 1979, Vol. 9, No. 4.

On Further Examination, Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline, College Entrance Examination Board, New York, 1977.

Rentz, Robert, "Testing and the Baccalaureate Degree," paper presented at ETS Invitational Conference on Testing, Washington, D.C., October 1978.

"Report to the Board of Higher Education on the Results of the New Jersey College Basic Skills Placement Testing, May 1, 1978–September 28, 1978," New Jersey Basic Skills Council, November 17, 1978.

Testimony before Senate Subcommittee on Education, Wednesday, March 28, 1979.

Identifying and Evaluating Major Elements In the Quality of Student Life

By Thomas R. Plough

College students constitute the most neglected, least understood element of the American academic community.

Frederick Rudolph
In *The American College and University: A History*

Identifying the Elements

Students will continue to enroll in institutions of higher education for reasons of career preparation, social status, and the availability of interesting fields of study. While most students are attracted to an institution by its academic programs and reputation, the quality of student life does influence the choice of college for some students, particularly those seeking academic programs not unique to any specific institution. Quality of student life is of greater interest in terms of its effect on student satisfaction and retention, and of alumni affection and support. Once enrolled, students tend to regard the efforts of the institution to maintain or upgrade the quality of student life as indicative of the degree to which the administration and faculty of an institution care about them. Seniors who have been involved in campus life experiences often view them as having more personal and practical value than their formal program of studies. Moreover, the opportunities provided on a campus for the expression of special avocational, athletic, and other talents is, for those students affected, nearly as important to them as their academic work. Knowledge and competence which can contribute to personal, social, and intellectual maturity can be substantially enhanced by the quality of student life on a campus. The satisfaction which comes from a high quality of student life often provides the strongest ongoing connection between students and their alma mater.

Elements that contribute to the quality of student life which extend beyond the important influence of the normal academic program include: involvement in special academic, athletic, social, cultural, and recreational programs afforded to students by institutions; caring relationships with peers and campus "mentors"; availability of needed services; access to facilities for recreation and socializing; reliable information sources about expectations for academic performance and social behavior; assistance in the maintenance of physical and emotional health and personal safety; courteous and prompt attention and replies to requests; and some attachment to a "community" within the institution or to the entire institution as a community.

The concept of student life is a difficult one to discuss because so few faculty and staff know very much about student experiences beyond the classroom and office. Any plans to upgrade or evaluate the quality of student life would be futile without an intentional program to inform members of the academic community about the realities of campus life. Special attention should be given to those positive and negative experiences which affect many students. Even in the absence of extensive student life research, ample literature is available which could target initial efforts towards understanding. James J. Jose sagely enumerated campus experiences which cause student dissatisfaction:

> abrupt and discourteous treatment in administrative and faculty offices
>
> failure to keep appointments with students
>
> uninformed advisors and counselors, particularly regarding institutional policies and procedures, coupled with an unwillingness to become knowledgeable except about their own specialized areas
>
> unwillingness on the part of faculty and administrators to take time to discuss student problems and opportunities
>
> the ease with which some administrators and faculty make students feel that they are an extra bother[1]

Each college has unique irritations in the student life area which could be ascertained and added to such a list. Positive restatements of each of Jose's items could serve as a list of campus experiences which lead to student satisfaction. Giving higher visibility to common campus experiences that positively affect students is often overlooked as an opportunity to improve student life. We tend to forget that emotional contagion can spread in either a positive or negative direction. While it is difficult for faculty to escape their appropriate role as critics, a conscious attempt to avoid reinforcing casual and poorly informed student gripes and criticisms can be very helpful. Whatever list of positive and negative experiences in student life were compiled, it would almost certainly reflect a special concern about the quality and quantity of student-faculty relationships. Indeed, many self-studies conducted by academic institutions address themselves to the issue of meaningful linkages between the academic program and student life. The "significant others" in fashioning such connections are the faculty. Almost seven years ago, a task force at Stanford University highlighted the need to increase the level and quality of faculty-student interaction as a means of revitalizing the student life component of undergraduate education. The study called for a rewards structure responsive to the

Thomas R. Plough is Associate Vice President for Student Affairs and Associate Professor at the Rochester Institute of Technology.

faculty role of mentoring, an increase in the amount of faculty evaluation of student work, some provision of more broadly informed academic advisors, structured assistance for students to learn how to utilize the resources of faculty and other support service staff and facilities, and active exploration of avenues for increasing the involvement of students both inside and out of the classroom and office.[2] A. W. Astin has concluded from recent longitudinal data that involvement with the faculty is a salient contributing factor in student satisfaction with the quality of their campus experience.

> Student-faculty interaction has a stronger relationship to student satisfaction with the college experience than any other involvement variable or, indeed, any other student or institutional characteristic. Students who interact frequently with faculty are more satisfied with all aspects of their institutional experiences, including student friendships, variety of courses, intellectual environment, and even administration of the institution. Finding ways to encourage greater personal contact between faculty and students might increase students' satisfaction with their college experience.[3]

It is curious that, given these and many other reminders of the importance of faculty-student interaction and student involvement in the affairs of a campus, so few long-range plans, statements of merit criteria, or plans of work contain explicit references to providing such opportunities. Of course, the Official College Bulletin always implies that these objectives are central to the mission of the institution. The removal of this kind of incongruence between marketing pieces and operating documents and procedures might constitute a modest effort aimed at improving the quality of student life.

Additional supporting evidence for initiatives which mesh the lives of students and faculty to improve the campus learning climate come from several sources. Ann Kieffer Bragg carefully outlined the socialization process in higher education and its implications for upgrading the quality of the undergraduate educational experience.[4] Her work thoroughly reinforces basic concepts which should govern the affairs of faculty and professional staff as teachers, mentors and role models. The socialization process, which involves the learning of both personal and professional skills, knowledge, attitudes, and values, emphasizes the importance of student involvement in academic and campus affairs; explicit expectations about performance and behavior, with plenty of feedback; responsiveness to students seeking evaluation of their work; and suggestions to students about the best applications for their unique talents and undeveloped competencies. Many students find their way to opportunities for extended evaluations and stimulating activities beyond the classroom. However, basic sociological theory and research advocate that a faculty engaged more fully in student life can maximize the potential for learning and motivation available from informal academic and campus experiences. This socialization perspective lends itself to the description which researcher C. Robert Pace would offer about teachers who influence students:

> This inventory of virtues defines pretty well what most people would regard as the ideal professor: a scholar and a researcher, a thorough teacher, constantly keeping his materials up-to-date, making his goals clear to students, demanding high standards of performance, provoking lively class discussions, and giving examinations which provide a genuine measure of students' understanding; at the same time, a friendly and mature individual, knowing his students by their first names, interested in their personal problems, going out of his way to be helpful, and reacting to hysteria or criticism with calm objectivity.[5]

Intelligent social critics have further argued the value of the faculty mentoring function from a human needs perspective. Suzanne Gordon, in her analysis of American life, has explored the need for companionship and the concomitant problem of loneliness:

> ... life becomes a bit like playing the stock market in a time of recession: one is less concerned with maximizing one's gains than with minimizing one's losses.[6]

Those who work closely with college students will recognize this scenario. Certainly, the undergraduate experience will be only partially satisfying to many students without some help on this problem. While much can be done by campus peers, especially mature upper-class students, adult perspective, insight, and credibility are needed. Without the continuity of caring and the special understanding provided by adult role models on the campus, the student world is too prone to embrace the extreme ends of the interpersonal spectrum. Gordon reminds us about those extremes and the human dilemma which results:

> We seem to have progressed past the world in which the potential of personal life was limited by consideration of others. But perhaps we have entered a world in which the self is paralyzed because nobody seems to care.[7]

Of course, faculty and professional staff are guilty of inappropriate swings back and forth on the educational practice continuum, which can accentuate problems in student life. Alison Bernstein has commented on this phenomenon in relation to our delivery of individualized instruction:

> A nightmare of mine goes like this: to correct the problems of colleges with massive enrollments, we create individual self-contained learning vehicles that propel students in and out of institutions without their having to touch or be touched by anyone save perhaps a lone learning mentor. We combat institutional size by enhancing an individual's feelings of separateness from others.[8]

This student loneliness issue and our possible response to it illustrate the relationship of quality of student life to educational practices on the campus.

Lewis Solmon and Nancy Ochsner have recently reported studies which should give us pause as we contemplate the improvement of student life.[9] Their data reveals that the college experience does not correlate with an appreciation for the arts, greater social responsibility and public service, or ongoing continuing education desires. In terms of use of leisure time, attending formal classes was rated at the bottom of the preference list. Reading for leisure was higher for college graduates than for noncollege graduates, but community service through organizational involvements ranked very low. Whatever college documents may describe or however

faculty and administrators might detail the benefits of higher education, it would seem that the behavior of adult role models in academic institutions is somehow reinforcing an aloofness and aloneness of personal and professional style and attitudes on the part of graduates. Cool, detached professional expertise may be alright in many settings but is clearly not very helpful for student development.

While this analysis places a heavy burden for quality of student life on the faculty, it is obvious that efforts of other professionals on campus are called for. There are many possible role models and mentors available on a campus. There is a fair amount of documentation that suggests the effectiveness of selected upper-class peer models for students in a variety of areas, including tutoring, advising and leadership training. Students are not, however, a substitute for adult role models. Just as the oldest sibling can contribute mightily to raising the younger children, it nevertheless remains a responsibility of the parents. The shared experience of student academic and social leaders working with faculty and staff to introduce, support and guide new students in various activities on the campus is often a peak experience for all involved. Such a program often significantly improves the lives of these upper-class students because of their interactions with faculty or staff supervisors.

Criteria for judging the quality of student life would include: high retention; solid alumni support; positive student attitudes expressed in campus media, seminars and surveys; good attendance at meetings and other events; absence of ongoing intragroup conflict; few occurrences of debilitating personal crises; and low incidence of vandalism and academic dishonesty.

If the research in sociology of work is correct, the academic leadership (top and middle management) of an institution can also influence the quality of student life in both direct and indirect ways. Honest communication with students about the bases for campus decision-making, combined with personal actions which indicate interest in student work and activity, can favorably affect the campus climate. Deans, department chairpersons and directors of support services occupy central positions from which the attitudes of caring can be communicated to students. Important role models and mentors for students are often found throughout the ranks of professional staff on a campus. The high regard which coaches, student activities staff, residence halls professionals and office secretaries receive from students is very impressive. While letters of reference are written by faculty and professional staff, it is often the secretaries who receive the little notes from graduates about how they're doing.

The extent to which a majority of faculty, upper-class students, academic management, professional staff and secretaries make serious attempts to interact with students in a caring manner; facilitate opportunities for students to engage in joint projects, programs, and activities with adults in the academic community; and clarify expectations for performance in academic work and campus behavior is directly related to the quality of student life. If this outreach is combined with a sound academic program of study, outlets for recreation and social activity, a handful of positive campus traditions, and accommodations to basic health and safety needs, continual enrollment and future affection for the college is assured on the part of a large number of students.

Criteria for Judging Quality

Criteria for judging the quality of student life would include: high retention; solid alumni support; positive student attitudes expressed in campus media, seminars and surveys; good attendance at meetings and other events; absence of ongoing intragroup conflict; few occurrences of debilitating personal crises; and low incidence of vandalism and academic dishonesty. Additional tip-offs might include the regularity with which faculty and staff are seen at campus events after five pm, and the presence of awards and rewards for faculty and staff who actively contribute to student-life activities. The ease with which most students could obtain letters of reference from one or two adult members of the academic community for knowledgeable appraisals of both academic performance and personal and professional competencies would also be a useful index of the quality of student life.

The overview of factors affecting student life presented thus far does not necessarily represent a substantial claim on incremental dollars. It does require some shifting of institutional priorities, attitudes and behavior. The costs associated with acting on this perspective are primarily psychological and social, not monetary. From a student perspective, this analysis should be quite ample to develop satisfactory quality in campus life. From a faculty perspective, and somewhat less from a professional-staff viewpoint, this call for active outreach is not without hidden sources of conflict. It is not necessarily true, regardless of the stated educational philosophy of an institution, that faculty and staff are rewarded for "mentoring" activities. This accountability to mission statement and follow-through on the allocation of scarce resources is the special obligation of the central leadership of a college. If the president, provost, deans or vice presidents do not explicitly and publicly affirm and reward activities which go beyond the traditional range of competent instruction, research, university service, and efficient administration, it would be futile to envision much advancement on the quality-of-student-life front. Requiring faculty and staff to do something, however important to student development and satisfaction, without prioritizing this assignment with some equity within the institutional reward structure would be hypocritical and counter-productive.

One approach to this matter could be targeted for the academic department. Each department chairperson

would be held responsible for achieving a dynamic balance in faculty staffing, which would satisfy several college objectives. Each departmental faculty would need to deploy a reasonable percentage of individuals who were (1) effective in monitoring students in academic and campus affairs, (2) creative in designing special ways to involve students in research or activities representing an extension of the classroom work, (3) gifted in classroom instruction, and (4) interested in and capable of linking up with campus programs and activities sponsored by professional staff or student leaders. A few faculty can practice all four roles and more. Most faculty cannot be all things to all students. Some faculty like to switch their emphasis from time to time. Some faculty would become more involved in student life if they were better informed about campus resources available to students. Every department, or larger unit in a small college, should be able to carry out this broadly based assignment. The institution should underwrite an academic leadership development program to better prepare chairpersons for this kind of responsibility. Such a seminar series might combine presentations and discussions on student characteristics, learning styles, nature of student life research and findings, the relationship between cognitive and affective development, the type of and availability of support services on the campus, methods for rewarding and evaluating faculty, and many other topics. Such a program series should be planned with ample input from the academic leaders who will participate in it.

Another matter which deserves some treatment in the quality-of-student-life area is the lack of a comparative perspective. Most students have no historical framework or interinstitutional analogies to help them evaluate the quality of their campus experiences. Left to their own, students can become disenchanted with campus life because they do not see constant upgrading. Every institution must provide some of this perspective through its orientation processes and its channels of communication. It is often necessary for the chief student development officer on a campus to assume a special obligation in this regard. Budget realities, based on institutional priorities and the primacy of the instructional programs, need to be fully articulated. There is also a real possibility that certain gains in the quality-of-life area will lead students to expect a further set of experiences. Leadership is required to moderate this Maslow-like progression when the realities of time, energy and funds make achievement unlikely. Honest communication to students on these matters, and genuine reminders of the current opportunity for an enriching student life are the answers to this complication. This is particularly true when evidence of advocacy for a student life program has already been demonstrated. Students often need to be helped to recall some of these advocacy circumstances, since short-term memory is a characteristic of transient student populations.

Sound academic programs are not sufficient for a high quality of student life. The role of the faculty must extend beyond providing excellent instruction. Admittedly, there are concerns which faculty and professional staff might raise about an "expanded" role. These concerns would probably focus on the nature of the reward structure and the priority activities which they must continue for reasons of potential advancement into other colleges or organizational settings or their own professional development and satisfaction. This is a challenge which must be addressed creatively and sensitively by the academic leadership at an institution.

A contention advanced in this paper is that the concept of mentorship, properly conceived and appropriately rewarded, can be a central and effective link between the world of the classroom and the social world of the student. While no argument should be made that contributions to the quality of student life are the top priority for all faculty and professional staff, or should be the sole merit criterion in an academic reward system, academic leadership must be accountable for seeing to it that mentors and other student life advisers and facilitators are available, active and affirmed. And while students themselves can make many improvements in the quality of their lives, there is a clear need for faculty and staff sponsorship, supervision, encouragement and guidance if continuity and quality are to be maintained over time.

Large incremental dollars are not required by this approach. However, if some incremental monies were to be appropriated for student life matters (faculty entertainment allowances, special student life research, academic leadership development programs), a good case could be made for the cost effectiveness of such expenditures. Certainly student retention and satisfaction data and its relation to alumni affection and support provide the basis for such a case defense. While a solid academic preparation is perhaps acceptable to students in return for their time and tuition, they are likely to see this kind of relationship as a *quid pro quo*. Upon graduation, this relationship between student and college ceases. To undergird a broader-based approach to student life, the academic leadership of a college must communicate frankly and regularly with students about the realities and priorities of the college. Special efforts (even white papers and presidential convocations) to update students on administrative plans and decisions which appear supportive as well as non-supportive to student life go a long way towards dispelling student mistrust, alienation and general dissatisfaction.

Evaluating the Elements

Each campus should determine the appropriate relationship between its formal academic programs and the quality of student life. The lay of the local turf will determine the priority given to efforts at involving students in special academic, social, recreational, and cultural programs. Out of the broad range of activities which contribute to student life, each college should select some to stress. Initially, an attempt must be made to determine whether a critical mass of constituents on the campus agree that quality of student life is an issue of genuine concern. A survey, interview assessment, or review of recently completed reports and campus media topics might indicate whether the quality of student life would find its way into a list of the top dozen issues to be assessed by the institution as it looks towards the next decade of operation. Often, certain trigger events can help to initiate such an overview. A new president,

provost, or vice president for student life might commission a task force to identify issues; the faculty senate might include this concern in their plan of work; a reaccreditation review might include this topic; the process of developing a new five- or ten-year plan might encompass this theme; the need to begin an endowment campaign may make concrete statements in this area extremely useful to development staff; or an inquisitive board of trustees or regents might inquire into the subject. It is surprising how many individuals will list quality of student life as a personal concern and an issue worthy of analysis. Obviously, if it is not a campus concern, there is no point in any active exploration of the topic. In this case, the normal processes and procedures of the student services organization on the campus can adequately facilitate ongoing efforts to maintain the current level of student life in consultation and cooperation with those students and faculty engaged in campus interaction and activities.

Once an inquiry has found this topic to be of interest to the campus, a study group with high credibility to all campus constituent groups should be given this task. Staff support, release from any other add-on assignments for study group members, and special consideration for a partial or full leave for the chairperson of this group are required. If possible all the work of this group should be completed in one academic year. It is best if the review is accomplished during a time frame when all students who were enrolled at the beginning of the study are still on campus at the time that tentative results are released. Various methods might be employed by this core investigating team, although a simple approach would be preferable to any complex arrangements. Public forums should be an integral part of this exploratory look at student life. This would be in addition to any review of literature, special studies, analysis of available data base, seminars, interviews and participant observations. It is important to plan these forums in consultation with academic advisors and professional staff who have strong associations with the student society in order to avoid the capture of such a study by vocal and assertive campus interest groups. This concern must be a central focus for the chairperson of the study team. The Stanford study, mentioned earlier in this paper, and a recent Middle States case study completed at Rochester Institute of Technology[10] offer two examples of reasonably successful endeavors in this regard. Ultimately, a report will be produced and should be widely circulated on campus for discussion and follow-through by the commissioning agent. In some cases, this study will be all that is required to set in motion the proper attention to features highlighted as conducive or inhibiting to quality of student life.

In other cases, the report of the campus study team might well recommend the need for more ample, reliable and valid data on the nature and quality of student life experiences. Such a data base would allow for year-to-year comparisons, and would provide planners with some ability to pick up subtle changes in professional orientations of faculty and staff or characteristics and needs of students. Both of these factors can interact over time to generate large incongruities between student life and institutional directions. The construction of a data base should follow the individual biases of a campus and its data base "managers." Some combination of the approaches associated with Mead, Gallop, or Astin, Pace and Sterns would be desirable. The use of structured observations and in-depth interviews, combined with timely campus polls and modified questionnaires with some open-ended categories should be carefully considered. The data battery should include a demographic profile done through the admissions process; an incoming student questionnaire on knowledge, attitudes, and behaviors, replicated in the second and final years of study; an exit interview or survey procedure; an alumni survey with some comparable items found on the other instruments; and for comparative perspective, occasional participation on a nationally standardized inventory. This type of data package will be useful for top management, long-range planners, and student development staff to ascertain the general relationship between student characteristics and campus experiences and the resulting quality of student life.

Often this base-line data portrait is neither sufficient nor stimulating to faculty who would like to see information on "their" students before any decisions on appropriate program responses are made by them and their academic leaders. Some data on student characteristics, experiences, and satisfactions tailored to departments, divisions, schools or colleges is needed. For this in-depth targeted data, case studies are the most illuminating and interesting "teaching" tools for the faculty. For an excellent example of the rich kind of data provided by case studies or diary-based studies, read "The Education of Kate Haracz" found in the May–June 1970 *Change* magazine. These case studies are normally done for the duration of a semester or academic year.

Beyond this kind of information, a longitudinal study of students at an institution is very useful. While this type of student life research involves a bit more expenditure of time and funds than those aforementioned, stronger functional relationships between specific campus life experiences and student satisfaction and persistence can be determined. If this element is added to the data collection system, it should be characterized by some open-ended items or structured interview questions. Straight questionnaire data miss far too much of the variety and originality of individual students' lives.

Polling mechanisms and theory are underutilized and misunderstood on most campuses. The value of polls for spot-checking "consumer" confidence in the student body on a variety of issues is quite high. Polls often reveal problems, concerns, and tendencies which motivate campus authorities to intervene in an effective preventative manner rather than on a reactive basis. Polls are also efficient for developing items and questions for other instruments contained in the data capture system. Polling faculty and professional staff can also contribute to a timely understanding of their attitudes and concerns. It is reasonable to hypothesize that there will be a positive correlation between the quality of faculty and staff life and the satisfaction and productivity of the students they teach.

If an academic institution is fortunate enough to have in place an academic advising structure which concentrates on mentorship rather than scheduling and course

selection, the advisers should be regularly interviewed or surveyed to identify their assessments of helpful and dissonant student life experiences. This would not be as useful a procedure if the advising structure were graduate student or peer based. If adviser input can be collaborated by similar reports from student service staff and students engaged in peer support programs, other data collection efforts can be modified to reflect these insightful leads. This kind of process also helps to insure the acceptability of student life research reports, which must be shared with faculty advisers. A data package which cannot be tied into an advisory system will be seriously weakened in terms of its influence on campus practices that affect quality of student life.

Generally, this approach to evaluation of student life need not be excessively expensive, especially if the effort is coordinated by a competent research director or team. Expenses in the student life research area get out of hand when evaluation is done *ad hoc*, represents duplication of effort, cannot be delivered in acceptable format to users of the data, is suspect in terms of design, sample, or analysis, cannot be individualized for major academic units, or resists modifications based on knowledgeable inputs.

Without a doubt, there are effective and efficient combinations and consolidations of the various research elements identified in this approach. Daniel L. Kegan has developed one such consolidation approach used at Hampshire College. Kegan's instrument "is a brief survey designed to include questions that will monitor key aspects of the quality of student life and effectiveness of the college."[11] His survey incorporates both standard forced-response items and some open-ended questions. Behavioral data (study time, adviser contacts), demographic information (class, financial aid), examples of learning (intellectual, social, physical), and process and personality items (acceptance in groups, enthusiasm) are included in the survey. All of these item categories can be modified when specific programs need to be assessed in terms of their functional impact on quality of student life. It is important to distribute and collect such surveys at approximately the same time in the academic year or at regular intervals throughout the year to avoid contamination by typical "depression" cycles of the academic year, weather patterns, vacation releases, and seasonal socialization patterns.

Summary

There are any number of ways to organize an institution to promote understanding and to assess the quality of student life. Whatever the structure of an academic institution, there is no way that the academic leadership can sit this one out. In terms of affirmation, accountability, rewards, and general decisionmaking, academic leaders symbolize and direct the priority given to quality-of-student-life concerns. It would be improbable that a high quality of student life would be found on any campus where the president, central officers, deans and chairpersons did not allocate personal time and institutional resources towards selected aspects of student life. Deferred humanism is a concept similar to deferred maintenance, and it is accompanied by equally painful long-range implications. Complete denial of any student life realities by academic leaders will lead to losses in enrollments and poor annual giving. Quality of student life moves students towards self-reliance and confidence rather than to the low-risk-client syndrome. The difference between caring for and pleasing students is basic, both in its philosophy and its manifestations. Advising and mentorship must be provided and rewarded. Competent individuals must be in charge of the student life research function, whether it be primarily a collection of knowledgeable observations from those who work closely with students or from a systematic data base. Regular and forthright communication about the priorities of academic and campus life issues to students must become traditional. Creative balances in the deploying of adult role models from the faculty and professional staff beyond the normal classroom and office setting must be achieved. Ultimately, quality of student life means a *responsiveness* to students as unique human beings. In this regard, small acts on the part of faculty and professional staff often add up to a caring community for students. The professor who takes time out to explain criticism of student work in a constructive way, the health service nurse who visits the hospitalized student, the president who sits down at a table in the cafeteria with a group of students, the computer services department that responds, in writing, in its newsletter to every comment and suggestion made to them by students for improving the service, the dean who writes a short note to students who appear in the campus paper as sponsors of a community service project, the professor of television production and the residence hall adviser who sponsor a group of students to develop a documentary on residential life, the ombudsperson, dean of undergraduate students, or director of student life research who clearly presents valid data which supports improvements in the lives of students, the student leaders who initiate a process to shadow incoming students for several weeks as they become acquainted with the college as an organization, and the secretary who takes the time to smile and track down the answer to a question of moment for a student—these are the agents of the university responsible for the quality of student life.

> The difference between a Duchess and a Charwoman lies not so much in the way they act as in the way they are treated. . .
>
> George Bernard Shaw

References

[1]James R. Jose, "Let's Talk About Keeping Those Students We Already Have," *Liberal Education*, December 1974, p. 509.

[2]*The Other Stanford*, A Report on the Relationships of Professors and Undergraduate Students at Stanford University, Chapter III, Stanford Workshops on Political and Social Issues, Stanford, California, 1973.

[3]Alexander W. Astin, *Four Critical Years*, Jossey-Bass, San Francisco, California, 1977, p. 223.

[4]Ann Kieffer Bragg, *The Socialization Process in Higher Education*, AAHE-ERIC/Higher Education Research Report No. 7, American Association for Higher Education, Washington, D.C., 1976.

[5]C. Robert Pace, "Perspectives on the Student and His College," *The Student in Higher Education*, American Council on Education, Washington, D.C., 1965, p. 88.

[6]Suzanne Gordon. *Lonely in America*, Simon and Schuster, New York, New York, 1976, p. 21.

[7]*Ibid.*

[8]Alison R. Bernstein, "How Big Is Too Big?", *Individualizing the System*, Jossey-Bass, San Francisco, California, 1976, p. 24.

[9]Lewis Solmon and Nancy Ochsner, "New Findings on the Effects of College," *Current Issues in Higher Education*, 1978 National Conference Series, American Association for Higher Education, Washington, D.C., 1978.

[10]Thomas R. Plough and Joanne Subtelny, Co-Chairpersons, *Quality of Student Life: Responsiveness in the 80's—A Case Study*, Rochester Institute of Technology, Rochester, New York, 1978.

[11]Daniel L. Kegan, "The Cycles Surveys," Longitudinal Indicators of the Quality of Student Life and a Framework for Evaluation and Administrative Experimentation, *Evaluation Quarterly*, Vol. 2, No. 2, May 1978, p. 294.

What Is the Most Effective Way To Handle Program Discontinuance?*

By Edward A. Dougherty

One of the most important things to keep in mind about program discontinuance is that it does not save money, in most cases.

Program discontinuance is a necessary adaptive mechanism for colleges and universities if they are to maintain flexibility and respond to new demands. While there is considerable resistance to cutting programs, some of that resistance can be overcome with foresight and planning, and we can learn a good deal from the experience of others who have already gone through the process. This study was initially undertaken to test the hypothesis that two variables would determine how an institution might respond to financial difficulties: one was the degree of financial difficulty and the other was the degree of control exercised by the institutional leaders. It was hypothesized that the more difficult the financial situation and the stronger the leadership, the more likely that an institution would resort to program discontinuance. What we found, in fact, was a good deal more complex and a good deal more interesting.

The sample for this study was drawn primarily from large public research universities in the United States, although two smaller regional campuses and one large private research university were also included. The characteristics and procedures of the universities studied are sufficiently similar to warrant viewing them together. The eleven institutions studied include:

Florida State University
Louisiana State University
State University of New York–Albany
University of California–Berkeley
University of California–Riverside
University of Michigan
University of Minnesota
University of Pennsylvania
University of Washington
University of Wisconsin–Eau Claire
University of Wisconsin–Madison

The selection of these universities was made on the basis of the fact that each institution (1) had reviewed and closed programs and/or (2) had gone through recent financial difficulties.

The author visited each of the campuses and interviewed key administrators and knowledgeable faculty. He saw academic vice presidents, graduate deans, deans, financial officers, institutional research staff, and faculty representatives from advisory committees or faculty senates. Data were collected from a variety of written documents such as financial statements, program review documents, and annual reports. In addition, representatives of state agencies were interviewed in every state where there are active coordinating or governing agencies.

Five areas of concern will be examined below: financial crisis, control and leadership, the role of the state agencies, the review process, and the process of program discontinuance.

Financial Crisis

One of the most important things to keep in mind about program discontinuance is that it does not save money, in most cases. At best, it reduces future expenditures, but it does not save money unless faculty are actually released. In most of the cases included in this study, the institutions did not want to break tenure contracts and they simply reassigned faculty to other duties on campus. Some funds may be saved by release of nontenured faculty, but even then proper notice is required and it may be a year or more before savings can be realized. One notable exception to this point is the State University of New York at Albany, which did cut programs and reduced faculty numbers; but even these funds were not released immediately. Many institutions and state agencies that undertake program discontinuance do so not with the intent of making immediate savings, but to improve efficiency in the long run.

If an institution wants to save money by cutting programs and reducing faculty, it should be careful not to cut programs that will reduce student-fee revenue in such a way as to cost the institution more than it is saving by reducing faculty. The University of Pennsylvania did a careful cost analysis before closing its program in Allied Health and found that it would be economical only if the student positions lost in the School could be picked up by other units on campus. Because the positions could be picked up by other schools and colleges at Penn without requiring the addition of new faculty positions, dollar savings could be realized in the long run. This type of

Edward A. Dougherty is Assistant to the Vice President for Academic Affairs at The University of Michigan

*Research supported in part by the Exxon Education Foundation.

cost-benefit analysis can be extremely important in making a decision about closing a program.

A second thing to keep in mind about program closure is that it cannot be done in the midst of severe financial crisis, in most cases. If a budget cut comes, it often comes in the middle of the year when funds are already committed. Such a crisis demands an immediate response in the short run, and most institutions do not have review and discontinuance procedures in place so as to allow for an immediate response. Again, Albany is the exception. It did have procedures in place and was able to put them into effect so that programs could be identified for discontinuance within about a five-week period.

When a financial crisis strikes a severe blow to an institution, it will tend to cut across the board rather than cut programs. Across-the-board cuts were made at Minnesota, Riverside, Washington, and Wisconsin and usually required taking faculty positions by attrition. Across-the-board cuts can be effective if they are not too severe nor too prolonged. They can weaken the whole fabric of quality in an institution if they are prolonged, however, and in some cases institutions will choose to cut out entire programs rather than cut across the board.

Third, while most institutions will make across-the-board cuts rather than program cuts in the face of severe financial crisis, some sense of financial urgency is usually necessary if the institution is to bring its faculty and students to accept program closure. One of the striking things about the institutions studied is how few closed programs simply on the basis of quality. Only Berkeley's closure of the School of Criminology was done without reference to financial considerations. All other institutions were prompted to begin a process of program review because of either a financial crisis or a "technical crisis," such as declining job markets for graduates of a particular program. In most institutions studied, the programs that were closed were called to the attention of the dean, vice president or coordinating agency because of some kind of crisis. The crisis was sometimes one of leadership or quality, but more often than not it was because of the loss of some source of external funds or the need for additional resources to maintain quality. For example, Michigan's Population Planning Program was cut when key sources of funds from outside the University were cut off and the department chairman asked the School of Public Health to pick up losses. The School of Public Health was going through its own financial difficulties and decided it would be better to cut the program out entirely rather than let it limp along. In many cases, financial or technical crises may not be controllable by the program or institution, but the program or institution should be careful to avoid crises of their own making if they wish to avoid program closure.

Another way to avoid closure is to build sufficient flexibility into faculty appointments so as to weather unexpected storms. The institutions that survived the depression of the mid-seventies by cutting only positions and not faculty had such flexibility. They tended to be either large enough to have sufficient turnover to make budget cuts simply by putting a freeze on hiring, or they had sufficient flexibility in non-tenured ranks to help them survive. The University of Wisconsin at Eau Claire has begun systematically to hire people outside tenure track positions in order to maintain flexibility in anticipation of future enrollment declines. The key question for institutions like Eau Claire is whether they will be able to gain sufficient flexibility in the face of 30-40 percent enrollment declines, or whether they will have to release tenured faculty as well. In some cases the answer to this question will determine whether or not they can survive. The University of California at Riverside has begun to identify a core number of faculty for each of its graduate programs. If the number falls below that minimum, then the administration will begin to look at whether to continue the program.

Program discontinuance will not save money in the short run, but it is, paradoxically, most often undertaken in the face of some financial difficulty. Careful financial analysis will be needed to understand what savings can be realized and to demonstrate to various constituencies the necessity of making such cuts. If one wants to avoid discontinuance, then a flexible staffing pattern may help the program or institution weather future storms.

Control and Leadership

The most important thing about control and leadership as it relates to program discontinuance is that there were virtually no cases in which programs were closed as a result of faculty actions. Faculty don't close programs, for a variety of reasons. They are concerned about the moral and legal obligations to their colleagues; they fear setting off a downward spiral of decline; and they are reluctant to make a final judgment of a colleague's quality. One dean said his faculty were suffering from "survivors' guilt"--that is, they would ask, "Why not me?" whenever a colleague was threatened by program closure.

The reluctance of faculty to make cuts in their own programs means that strong leadership is called for to pursue what may be very unpopular decisions. The job of the administrator in program discontinuance is primarily one of convincing others of the rightness of his or her judgment. It is largely a job of convincing all constituencies that the institution will be better off without program "X."

What are the tools with which the administrator has to work? The first is the review document. This will be discussed in detail in an upcoming section, but it should be noted here that the review document will be the most important item available to the person whose job it is to implement a discontinuance. The purpose of the review document is to expand the pool of expertise to include more than just the dean or vice president or president recommending discontinuance. It must include judgments from trusted peers who are well respected and whose conclusions follow in logical sequence. The discontinuance of the Berkeley Criminology program is an example of a well-documented review. The Sindler Committee did an extremely thorough job in the six-month period of deliberations. It met with all interested parties and drew its conclusions in such a way as to leave no room for doubt about the quality of the program. Committee members identified the core problem of the School as not being able to meet its own stated goals of providing both academic and professional training. After carefully documenting this failure, they then asked if

criminology should continue in some form as an academic discipline on the Berkeley campus. They examined the undergraduate and graduate programs and concluded that it should be continued, but on a much smaller scale and with a very different direction. They concluded by suggesting that such a program might be located in the Law School, and their recommendations were eventually implemented. The report was thorough and clear enough to allow the graduate dean and chancellor to recommend with confidence the discontinuance of the School.

With the help of a well-documented review, the academic administrator can turn his or her energy to the task of eliciting support for a proposed discontinuance. Support must come from two groups—the faculty and the board that has final authority for eliminating a program. Faculty must have full opportunity to express their views on a proposed closure through open meetings, forums, debates and discussions. Students, too, must have the same opportunity to present their views. At Michigan we met with faculty in the program and faculty outside the program at their meetings, and we held open hearings for them to present their views to us. Administrators must educate the board of trustees or other groups on the border between the university and the outside community. Thomas Langfitt at the University of Pennsylvania did an especially good job of working with the health affairs subcommittee of Penn's Board before making a final recommendation to the full Board.

The credibility of the administration closing a program is an important factor in leadership, although it is a very difficult quality to measure or discuss rationally. If faculty or board members do not trust the administration, the administration's job is much more difficult. It will be accused of heavy-handedness, favoritism or politics. Innumerable secondary issues will be thrown into the discussion and opposition will arise at every turn. If, on the other hand, the faculty believe in the fairness and good judgment of the administration, then they will be open to discussion of the issues. The central administration of the University of North Carolina (not originally included in this study, but recently visited by the author) is a good example of the importance of credibility. The administration has been able to implement a process of program review and discontinuance, in part because of the trust faculty and campus administrators have in President William Friday.

Lack of leadership clearly has a negative effect on an institution's capability to implement program discontinuance. At several of the campuses studied the turnover of key administrative positions contributed to the lack of action in programmatic areas. In one institution the president, academic vice president and graduate dean all left within 18 months. These changes, combined with the lack of clear budgetary authority in any one of the vice presidential areas, mean that no programs have been closed in spite of financial strain. It may be no program closure was called for in this situation; but the existence of rapid administrative turnover and no program discontinuance was repeated on other campuses.

Another pattern of leadership that is repeated on several campuses is the role of the graduate dean in the process of discontinuance. For example, at UC-Berkeley, Florida State, Louisiana State and UC-Riverside, the graduate dean controls the flow of admission into graduate programs. Second, the graduate dean is often vested with authority to review programs and can spot potential areas of weakness. Third, the graduate dean often has the credibility to review and discontinue programs because she or he is seen as protecting centers of excellence in the institution. All of these factors can be important in the process of review and discontinuance.

The Role of State Agencies

Robert Berdahl suggests three important roles for the statewide agency in the area of program review and discontinuance: to do reviews for institutions that are currently not doing them, to assure that the quality of review is maintained, and to assure that evaluations are placed in the broader context of statewide, regional, and national needs. Berdahl estimates that in 1975 eleven of the twenty-eight coordinating boards were engaged in some form of program review, and two others were in the process of developing review procedures.[1] It is likely that today there are even more boards doing more reviews, especially if the nineteen consolidated governing boards are included in the count. One finds generally an overall trend toward more formal systems of coordination and program review at the state level, and the research universities studied in this paper are no exception. California, Florida, Louisiana, New York, Washington and Wisconsin have all undertaken some form of statewide program review since 1970. The Universities of Michigan and Minnesota operate under a constitutional provision that gives them constitutional autonomy, but Michigan does coordinate new programs in the state through the Michigan Council of College and University Academic Officers, and Minnesota cooperates with reviews conducted by the Minnesota Higher Education Coordinating Commission.

The primary purpose for these statewide reviews is to provide the state, regional or national perspectives on program planning. For example, the lateral reviews done in states such as California, Florida, Louisiana, New York, Washington and Wisconsin normally examine all programs in one discipline or professional area to determine the needs of the state for graduates in that area. Review teams may then look at the quality of existing programs and make recommendations about the future needs of the system.

At the same time state agencies are doing reviews, institutions are doing very much the same kind of reviews on a local level. All the institutions in this study conduct reviews on their own, the only exception being Louisiana State, which stopped its review process when the state agency began theirs because of the extreme time demand placed on the institution by the external review process. There may be some smaller institutions that do not do reviews or that do not do high-quality reviews, but they would be the exception. The review process done at the state level may duplicate what goes on at the campus level. One should not belittle, however, the state, regional or national perspective added by the state agency review. For example, in Florida a review of

nursing pointed out national trends in nursing education that were not being followed within the state. Large amounts of money and effort were being spent on the preparation of nurses to deliver secondary care, but the trends indicate a need for more nurses prepared to deliver care at primary and tertiary levels. With that information, the state agency and the institutions are beginning to plan for the appropriate number of nurses to be trained for needed levels of practice.[2]

In addition to providing a broader perspective to the review process, it is sometimes pointed out that state agencies are interested more in a "summative" review process, as opposed to a "formative" process. (The summative review is aimed at making a decision about whether to continue or discontinue a program, while a formative review is interested in making suggestions about how a program can be improved.) Looking closely at the relationship between internal and external reviews and the kind of decisions that result from them, one does not find this distinction to hold up, at least in terms of relating reviews to program closure. Figure 1 shows that four institutions (UC-Berkeley, Michigan, Penn and UC-Riverside) are primarily involved with internal reviews and have closed programs without any help from the state agency.

Figure 1. Internal and External Reviews Related to Program Closure

	Reviews Internal Only	Reviews Internal and External	
Resulting in Program Closure	Berkeley Michigan Penn Riverside	Albany ..... Albany Eau Claire Florida ..... Florida Louisiana Washington ... Washington	
Not Resulting in Program Closure	Minnesota	Louisiana Madison	

In states where both internal and external reviews have taken place, programs have been closed after internal reviews as often as after external reviews. Only in Louisiana did the state agency, after external review, close a program that would not have been closed by the institution itself. Pressure may have been created to move the institutions to review and close programs because of the existence of state agency review procedures, but one cannot argue that institutions do not close programs on their own. Institutions without state agency review do close programs and institutions with such reviews close programs. The presence or absence of a state agency and statewide review procedure is not a decisive factor in determining the presence or absence of program closure.

Furthermore, in states and institutions where no formal program closure has occurred, there are continuing "program adjustments" that change programs. These changes serve as a form of program closure over the long run. They are, certainly, adaptive mechanisms which are frequently used by institutions. For example, the University of Minnesota and the University of Wisconsin at Madison continually review the quality of programs. Neither institution has closed programs, but adjustments have been made to reflect changing demand and quality.

The primary function of state-level review is to provide a broader perspective on the review process, but questions can be raised about the efficiency of duplicating reviews in order to provide that perspective. If there were a greater level of trust between the state agencies and the institutions, then more efficient ways might be found for getting that job done. In some cases the level of trust is good, such as in Wisconsin where the major research universities are allowed to do their own reviews. In other states there is little trust. For example, in one state the state agency representatives do not feel that any objective evaluation can be done by faculty. Even outside reviewers are suspect because it is believed they can easily represent the position of colleagues on the campus under review. This particular state agency relies solely on quantitative data, saying that a high-quality program will have positive quantitative indicators. One result of this approach is a good deal of conflict between the state agency and the campuses when the state agency recommends program closure. When the institution and state agency take up battle stations, little of the broader state, regional, or national perspective comes through to the institution. When the state agency wins individual battles, it can impose its will on the institution, but the institution, in turn, does everything it can to subvert the intention of the state agency, and little progress is made. A more cooperative approach could lead to gains for both the state agency and the institution.

Rather than offering a "political" or a "collegial" model for understanding the relations between state agencies and institutions, Martorana and Kuhns offer a "theatrical" model. They note that the present state of activities can be likened to "a theatrical production in its early rehearsal stages: the roles have been more than tentatively assigned, but the actors are still jockeying for position and attempting to upstage one another."[3]

The Review Process

Whether or not the state agency and the institutions cooperate, the single most important factor in program discontinuance is a flexible and fair review process. In the words of Barak and Berdahl:

> When participants viewed the process as flexible and fair, they then tended to perceive the overall review more favorably despite any identified deficiencies in the selection of programs reviewed or the criteria used to review them. The latter two, if deficit, were often viewed as susceptible to later improvement through a fair and flexible process.[4]

One might add that when the process is not viewed as flexible and fair, no amount of tinkering later on will help much.

There are four basic questions that must be addressed concerning program review: Who has the authority to review? What programs will be reviewed? What criteria

will be used? And, what procedures are most appropriate for the situation?

First, the question of authority: Who has the authority to review programs? The faculty? The campus administration? The graduate school? The state agency? Looking at Figure 2, one sees that the answer is "yes" to all options. No one pattern prevails. Looking at internal authority to review programs, one sees that the graduate school often has authority to review graduate programs, but that pattern is not universal. In some cases graduate and undergraduate programs are reviewed together by a special task force or review committee that may be part of the faculty senate. In other cases, the authority for review may rest with the deans and/or vice presidents for academic affairs. In some cases the authority to review programs shifts from one group to another on the same campus, depending on the circumstances. Most campuses do have some informal agreement about the authority to review programs, but as questions of review inevitably lead to questions of continuance and discontinuance of programs, those informal agreements may need to be more formalized.

The question of external authority to review programs is more complex. Berdahl points out that in theory all system-wide governing boards, such as the boards of regents in California, Florida and Wisconsin, have authority to review programs.[5] These boards, however, may delegate that authority to the individual campuses, as happens in California and Wisconsin, retaining for themselves authority to review any programs in a given area as the need arises. For example, Wisconsin allowed its major graduate campuses at Madison and Milwaukee to review their own masters and doctoral degree programs but reserved the authority for system-wide audit of masters and specialist programs on other campuses. Wisconsin also does lateral reviews in areas such as classics or computer science to determine overall state needs. At the same time, more detailed program reviews are done internally on every campus.

The authority for program review is not always clear in the state coordinating agencies (as opposed to governing boards). Coordinating boards vary widely in their power to review existing programs, ranging from no action to power to recommend program discontinuance. The University of Minnesota, for example, cooperates voluntarily with the Minnesota Higher Education Coordinating Board, even though the University has constitutional autonomy and the Coordinating Board has no authority over the University. On the other hand, the State Education Department in New York, which functions as a coordinating body for all of education, public and private, has authority to review and "deregister" programs. Again, the trend is toward the assumption by state agencies of more authority for review, and in some cases for discontinuance. There is also a trend toward formalizing the review process within those systems.

The second question is: What programs will be reviewed? Berdahl points out the need to select those few programs that can be evaluated in depth within the capacity of the group doing the reviews. To review on a more massive basis will result in either great superficiality or an overloaded board and/or budget.[6] The sample of institutions in this study show both extremes of superficiality and overwork. In the State of Washington virtually all graduate-PhD programs were reviewed at once. An initial audit of programs identified programs with low-degree productivity. A second audit identified some 300 duplicate programs that were subsequently all reviewed at once. Only quantitative indicators were used in this review, yet it took nearly three years to complete. Many faculty criticized the process as being superficial and unproductive.

A more typical pattern is to use the quantitative indicators to select programs for review, allowing for a few comprehensive reviews. As one dean said, a college executive committee or graduate council simply cannot absorb data from more than six or eight programs a year, especially if they are involved in other ongoing business. At both the state agency and the institutional levels some "trip-wire" is needed to trigger the review mechanism, or some regular cycle of review is necessary. One should not assume, however, that only negative indicators will trigger the review mechanism. Several institutions have found it advisable to review one or two programs of high quality, having positive indicators, along with one or two of suspected low quality, having negative indicators. These indicators will most often be indices like degree productivity, faculty productivity and student credit hours.

Other state agencies and institutions will review programs in logical clusters. For example, the Florida Board reviews all programs of one discipline across the state. Some institutions select all programs in related areas, such as natural science, social science or humanities, for review in a given period. Institutions like Michigan and Penn review programs in conjunction with the appointment of a dean or department chairman, which allows for the option of changing personnel if the program is weak, or selecting personnel with particular strengths to match the needs of the program. Other institutions, like Berkeley, review six to eight programs a year, some good and some weak, selected on the basis of logical groupings. Ad-hoc reviews will also be conducted as issues arise. It was one of these ad-hoc reviews that eventually led to the discontinuance of the School of Criminology at Berkeley.

The third question is about criteria. What criteria will be used to evaluate programs? Mingle identifies two distinct types of reviews that vary by the criteria used for evaluating programs.[7] One type of review is quantitative, concerned with such things as degree productivity, faculty work loads, and student credit hours. The other type of review is qualitative, dealing with a broader set of criteria such as the quality of faculty as judged by peers, the quality of students, and the nature of the curriculum. In most cases the campuses and the state agencies combine these criteria in some fashion. The state agencies are more likely to look at quantitative factors such as need, degree productivity, and work loads, but these quantitative factors may also be used by the internal review committees either to trigger a review or to be used along with a more qualitative evaluation. We have already cited as an example the Washington Council on Postsecondary Education, which used only quantitative data. By way of contrast we might cite the Berkeley review of the School of Criminology, which

Figure 2. Program Review Procedures

	Who Has Authority To Review	Selection of Programs for Review	Criteria*	Process	Outcome
Albany	*external*— State Education Department	quantitative indicators 3-4 reviews per year	quality and need	lateral reviews, external evaluators	programs cut
	internal— Graduate office '75-select Comm. '76-task force	all departments reviewed on 7-year cycle	quality	internal evaluators	programs cut
California— Berkeley	*external*— California System	ad hoc, as issues arise	quality and need	lateral reviews, external/ internal evaluators	no cuts, aid in planning
	internal— Graduate school	ad hoc as issues arise plus 6-8 graduate reviews per year	quality	internal evaluators, external consultants	mostly formative, one school cut
Florida State	*external*— Florida Board of Regents	quantitative indicators	quality and need	lateral reviews, external evaluators	one program cut— aid in planning
	internal— Graduate Council	all graduate programs reviewed on 6-year cycle	quality	internal evaluation, outside consultants may be used	primarily for program improvement but one school cut
Louisiana State	*external*— Board of Regents	all PhD programs reviewed	quality	external evaluators	programs cut
	internal— Graduate School	suspended when external process began			
Michigan	*internal*— Vice President for Academic Affairs and schools and colleges	all programs reviewed at end of term, appts. of dean or dept. head	quality	varied	primarily formative but one program cut
Minnesota	*internal*— Graduate School and schools and colleges	all programs reviewed on regular cycle	quality	external evaluators	formative—no programs cut
Pennsylvania	*internal*— Vice presidents and schools and colleges	ad hoc as issues arise plus 6-year reviews coordinated with deanship appointments	quality and cost	internal evaluators	one school cut

*See individual case studies for details.

	Authority	Selection	Criteria	Process	Outcome
California–Riverside	*external*–California System	ad hoc as issues arise	quality and need	lateral reviewers external/internal evaluators	no cuts–aid in planning
	internal–Graduate Council	all graduate programs reviewed in order of establishment	quality and technical factors	internal evaluators	several programs cut faculty lost by attrition
	Committee on Educational Policy	review of undergraduate programs abandoned			
Washington	*external*–Council on Postsecondary Education	quantitative indicators	quantitative indicators	council reviewed day	programs cut
	internal–graduate schools and schools and colleges	ad hoc 4-5 per year	quality and need	internal evaluators	one program consolidated
Wisconsin-Eau Claire	*external*–Systemwide Administration	quantitative indicators–MA & Specialist degree programs plus ad hoc as issues arise	quality and need	lateral reviews–internal and external evaluators	planning aid–no programs closed at Eau Claire
	internal–Academic Policies Committee	quantitative indicators 4-year cycle for undergrad. programs	quality	internal evaluators	some programs cut
	Graduate Council	graduate programs on regular cycle	quality	internal evaluators	some programs cut
Wisconsin-Madison	*external*–	ad hoc as issues arise	quality and need	lateral reviews internal/external evaluators	planning aid–no cuts at Madison
	internal–Academic Affairs plus schools and colleges	quantitative indicators 10-year cycle in LLS	quality	internal evaluators	formative

looked only at the quality of the program. The review committee looked only at qualitative issues such as how well the program was fulfilling its mission and the quality of its offerings. In fact, the quantitative indicators for the program were quite positive, but the qualitative factors left much to be desired in the eyes of the review committee.

The Education Commission of the States outlined 10 criteria that are frequently used by various states in evaluating programs. These criteria combine the quantitative and qualitative and reflect the concern of many state agencies. They are:

(1) the number of graduates in each of the last five years
(2) enrollment
(3) size of classes and cost per course
(4) cost per program graduate
(5) faculty work load
(6) program quality as reflected in national rankings, faculty qualifications, positions of graduates
(7) total production of graduates
(8) economies or improvements in quality if consolidated or eliminated
(9) student interests demand
(10) appropriateness of program to institutional mission[8]

The National Board emphasizes that single measures of quality should not be applied to very diverse programs, but rather, multiple indicators of quality should be developed. They also say that statewide planners should resist the temptation to apply simplistic formulas to doctoral programs, such as the elimination of any programs with fewer than two graduates per year. Such measures might be used as signals to flag programs in need of further review, but should not be used as a sole basis for elimination.

Russo and Brown focus on five criteria that are more typical of those used by internal review committees. They are:

(1) the quality of instruction and learning
(2) the quality of faculty
(3) centrality of program's contribution to the institution and other universities
(4) program value and/or uniqueness
(5) potential

They propose careful data collection and analysis that would allow evaluators to rank programs numerically by these criteria. Programs with different objectives could weigh the criteria used to judge success in meeting program objectives.[9] This numerical approach has the advantage of lending an aura of objectivity to program evaluation, but it does not hide the fact that in the final analysis the judgments of quality are still judgments of individual evaluators.

Finally, the question of process: What is the most appropriate methodology for program review? Once again, no single pattern of evaluation is universal. At best the methodology selected is the one that is most appropriate to the circumstances. At worst, the method is selected on the basis of political convenience. Don Gardner offers five definitions of evaluation based on methodological considerations. They are:

(1) evaluation as professional judgment, with professionals lending expert opinion on programs
(2) evaluation as measurement, which uses some formalized instruments of evaluation
(3) evaluation as the assessment of congruence between performance and objectives, which compares measured performance against specified goals
(4) decision-oriented evaluation, which examines performance to continuously assess programs
(5) goal-free evaluation, which examines performance regardless of goals.[10]

The most common method of evaluation used in the universities included in this study is evaluation based on professional judgment. The professionals may be from outside the program or outside the institution. In some cases they may include one or two members of the program under review. At Michigan we have found one effective method of evaluation is to use three evaluators from outside the University: one from the area being studied, one from a related area, and one from an unrelated area. The purpose of including an unrelated evaluator is to provide the opportunity for someone to raise the question of whether the review is worth doing at all.

Using peer evaluators seems to work best for institutionally oriented reviews that are formative in nature—that is, aimed at identifying areas of program improvement. This pattern can also be used for summative reviews, aimed at a decision to continue or discontinue a program. With the exception of decisions made by Washington's Council on Postsecondary Education, no program has been discontinued without some form of peer review. In some cases the institution's evaluators are internal and in some cases external; the important thing seems to be the rendering of some professional judgment on the quality and/or effectiveness of the program.

The Process of Program Discontinuance

With all that has gone before it, the actual process of program discontinuance may seem rather anticlimactic. At some point a dean, a vice president, or a board must say that no more students will be admitted to program "X" and that faculty will either be released or given other duties.

Again, the question of authority for program discontinuance must be addressed. Usually the governing board has the final authority to make such decisions, but in many cases the real battles have already been fought by the time the decision gets to a board. The board's role, however, may be an important one in terms of fending off external constituents, who will inevitably claim that no university worthy of the name can exist without program "X." Clear lines of authority should not be overlooked in the process of program discontinuance. Because the process will inevitably be political in nature, the board will want to listen to one or more of those external constituents and may reverse the recommendation to discontinue a given program. For this reason, the

education of the board to the realities of the situation will be as important as the education of the faculty and students involved. Thomas Langfitt at Penn recognized this educational task and spent considerable time with a committee of the Board of Trustees going over the review document that led to the discontinuance of the School of Allied Medical Professions.

What kinds of programs get closed? A sample of eleven institutions is not one from which conclusions can be drawn, but there are some interesting trends. In the major research universities the programs that are most likely to be discontinued are those outside commonly accepted definitions of the university. They often are professional training programs that do little to contribute to the research mission of the university, as ill-defined and imprecise as that mission may be. One of the reasons program closure is so difficult is because only programs that clearly go beyond the definition of the institution get cut out, under normal circumstances.

Being "on the outs" may mean that a program is not meeting quality standards or economic/technical standards. It may also mean that the institution simply does not want to include a given program within its scope of activities. For example, Berkeley decided that the orientation of its School of Criminology was no longer appropriate to the definition of a professional school as defined by the Berkeley review team. In the same way, Penn decided that its School of Allied Medical Professions, although strong among its peers, was no longer appropriate to the definition of Penn as a major research university. The School offered only undergraduate courses and did little research. It was therefore ruled to be outside the definition of the university when demands were made for its increased support.

In some cases the state agency helps define the appropriateness of a given program to the mission of either the institution itself or the state educational system as a whole. For example, New York's State Education Department defined history and English graduate programs as no longer appropriate to the needs of the state. Louisiana's Board of Regents said the graduate program in political science at Louisiana State University did not meet the appropriate definition of quality. In both cases the campus representatives argued strongly for the maintenance of these programs as essential to the definition of their institutions. The programs under discussion were seen as essential to the "core" of the university.

When a program no longer contributes to the mission of the institution or state, it becomes vulnerable to discontinuance. That mission may be defined in quantitative, qualitative or programmatic terms. It may be defined loosely or arbitrarily by one individual—for example, a president who wishes to give his or her institution a clear focus; or it may be defined by consensus from a large number of constituencies. The mission may never be very precise in the face of program closure. Depending on the circumstances, the institution may come to the decision that program "X" is no longer within the bounds of the mission of that institution.

Often the demand for increased support or some other catalytic event such as the search for new leadership raises the question of whether the program contributes sufficiently to the mission of the broader unit. For example, at Penn, the School of Allied Medical Professions asked for increased support from the central administration in order to qualify for federal funds. At Michigan the withdrawal of outside funds from the Population Planning Program and the request of the department for increased support from the School of Public Health prompted the Dean and Executive Committee to look more closely at the program's contribution to the School's mission. Eau Claire tends to be more sensitive to market changes for its graduates, and it defines its mission more in terms of market demands for its graduates. Hence, programs phased out there were ones with declining student demand. Similar considerations operated at Florida State when it eliminated its School of Engineering Science, and when market conditions changed, the administration regretted the closure. Albany, on the other hand, was forced to redefine and narrow its mission not because of changes in the market for its graduates, but because of severe cuts in state appropriations. Once again, we are reminded that it is difficult, if not impossible, to separate financial and quality or mission-oriented concerns. Dollars will often define quality and mission, and a dollar crisis may be the catalyst leading to redefinitions of quality and mission. It is important to note, however, that programs do not get cut only because of a dollar crisis. Qualitative and mission-oriented concerns may be addressed outside the sphere of financial difficulties, as seen at Berkeley and Louisiana.

Another factor that is important in determining whether a program will be closed is "opportunity." In several of the cases studied there was an opportunity for moving tenured faculty without violating the tenure obligation. At the University of Washington, for example, the Home Economics Department could close several of its programs because of the opportunity to move tenured faculty. Several faculty had voluntarily left and several retirements were possible. Other faculty could be shifted to new assignments. There were also a small number of tenured faculty involved in the Berkeley-Criminology closing. Out of more than 20 faculty in the program, only five were tenured. All five were shifted to other schools and colleges on campus. The University of Pennsylvania School of Allied Medical Professions also had few tenured faculty. Most will find positions outside Penn, and one will move to another position within the Medical School. Often, without the opportunity to deal with tenured faculty, no program will be closed. However, there are circumstances under which program closure will take place even without that opportunity, such as at Albany where the financial conditions of the state required rather radical definition of the institution's scope and mission.

Finally, programs that are discontinued tend to be isolated from other programs and departments, some physically. It might be more accurate to say that those programs that are not discontinued have a large number of supporters who say, at the critical moment, "We cannot get along without program 'Y'." Program "Y" provides some vital services that others feel are essential. Program "X," on the other hand, finds too few to lend support when the time comes to consider its continu-

ance. At Michigan the Population Planning Program was providing good service to foreign nationals, but departments in the School of Public Health or elsewhere in the University did not say that the services of the program were essential to *their* operation. By contrast, several units, including the University Hospital at Michigan, said that the Program in Speech and Hearing was essential to their operations and should be continued—if not as a part of the Medical School, then as part of another unit in the University. The Program was shifted to the Special Education Program in the School of Education.

Recommendations

What can be done to avoid discontinuance?

(1) Set priorities based on a clear understanding of the institution's mission.
(2) Develop curricular and faculty relations with other units.
(3) Expand functions within program guidelines and don't overspecialize curriculum.
(4) Do long-range planning.
(5) Don't let quality deteriorate.
(6) Maintain strong leadership.
(7) Avoid financial difficulties of your own making.
(8) Keep an eye on student numbers, both enrollment and placement of graduates.
(9) Build in a sufficient number of flexible, nontenured appointments to weather unexpected storms.

What should be done if closure is desirable or unavailable?

(1) Establish a line of authority for program closure.
(2) Establish agreed-upon review procedures.
(3) Establish procedures for making priority decisions.
(4) Establish procedures for giving notice to or relocating faculty and staff.
(5) Provide for input from faculty and students and for handling their needs, legal and otherwise.
(6) If programs are cut to save money, cut programs where real dollar savings can be made, not ones that produce more revenue than expenditures.

Footnotes

[1]Robert Berdahl, "Criteria and Strategies for Program Discontinuance and Institutional Closure," (paper presented for the Kellogg In-Service Training Seminar, BHEEO Annual Conference, New Orleans, LA., July 30, 1975).

[2]*State University System of Florida Nursing Program Review: Consultants Report* (Tallahassee, FL.: Board of Regents, 1977), p. 31.

[3]S. V. Martorana & Eileen Kuhns, "Academic Programming," AAHE/ERIC Research Currents (September 1978), p. 11.

[4]Robert J. Barak and Robert O. Berdahl, "State-Level Academic Program Review" (Report #107, prepared for the In-Service Education Program, Education Commission of the States, February 1978), p. 18.

[5]Berdahl, "Criteria and Strategies," 1975, pp. 5-6.

[6]*Ibid.*, p. 10.

[7]James R. Mingle, "Influencing Academic Outcomes: The Power and Impact of Statewide Program Reviews." In C. Fincher, W. T. Furniss, J. R. Mingle, and B. S. Spence, *The Closing System of Academic Employment.* Atlanta, GA: Southern Regional Education Board, 1978, pp. 51-68.

[8]National Board of Graduate Education, *Doctoral Manpower Forecasts and Policy*, No. 2 (Washington, D.C.: N.B.G.E., 1973), p.20.

[9]J. Robert Russo and David G. Brown, "A Model for Internal Program Review," *College and University*, Vol. LII, No. 3 (Spring 1977), pp. 288-298.

[10]Don E. Gardner, "Five Evaluation Frameworks: Implications for Decision-Making in Higher Education, *Journal of Higher Education*, Vol. XLVIII, No. 5 (Sept/Oct, 1977), pp. 571-593.

How Do You Know a Good College When You Stumble Over One?*

By John Harris

The answer to the question posed in the title of this paper might seem simple and obvious. Down my way, we like Tennessee Walking Horses. You would think anybody could tell if a horse walks "good"; actually, almost anyone can. For instance, you can take someone who has never seen a Walking Horse to a horse show, and while he is eating country ham and sipping Pepsi, he is likely to pick the same horses picked by the experienced judges. His untouched instincts about form and movement are often reliable. Similarly, many people without academic experience or training will choose as outstanding the same institutions that individuals with years of academic training and experience will choose.

What, then, do experience and knowledge contribute to decisionmaking? The critical contribution by the expert is his or her explanation of a decision; the layman's judgment may be reliable in that it is reasonably consistent, and valid because he often agrees with the expert, but he cannot "particularize" his judgments the way the expert can.

Knowing why a college is "good" is important if you wish to do any of the following:

(1) Explain or defend a judgment
(2) Diagnose institutional weaknesses in order to treat them
(3) Predict what will make a "good" college before one is ever developed.

Assessing Educational Quality

Bowen[1] has called for the use of "criticism" as the foundation for assessing educational quality, as distinct from the more numeric approaches characteristic of current forms of accountability. Such criticism must, Bowen says, take results into account.

> The study of outcomes should avoid the common confusion of inputs and outputs. Most evaluation of institutions is conducted in terms of such variables as faculty-student ratios, number of PhD's on the faculty, size of endowment, current expenditures per student, SAT scores, number of library books, range of facilities and equipment, value of physical plant, extracurricular programs, etc. These factors are all inputs and it is by no means established that there is any systematic and positive correlation between these inputs and the true outcomes defined in terms of the personal development of students.[2]

John Harris is Professor of Education at Middle Tennessee State University.

My limited experience with regional, institutional accreditation convinces me it is overly dependent on inputs in assessing institutional quality. My participation in the recent COPA-sponsored study of the accreditation of nontraditional higher education further convinced me. I found that visiting teams rarely attempt to look at educational results.[3] Troutt's unpublished dissertation, "The Quality Assurance Function of Regional Accreditation," confirms that the stated policies and criteria for accreditation take little direct account of results.[4]

It is time to look in the horse's mouth and count the teeth. What do students actually know? What can they do? Do they know and can they do what people in general think degree-holders should know and do?

A reemphasis on results in appraising quality would have several advantages for American higher education. I have tried to deal with these general advantages in other papers[5] as well as with their consequent implications for state coordination, budgeting, and accreditation.[6] My general thesis is that the relationship between the "processes" and "results" of education is not strong enough for us to assume the desired "results" will occur as expected simply because the approved "processes" are in place. That is, to account for student progress toward degrees by time units (students' credit hours) falsely assumes that time of studenthood is highly correlated with student achievement. There is also the assumption that if the means of learning are available—lectures, laboratories, libraries, etc.—students will and do reach certain achievements.

It is time to look in the horse's mouth and count the teeth. What do students actually know? What can they do? Do they know and can they do what people in general think degree-holders should know and do?

These questions raise the larger question of whether our educational means and timetables are producing the expected results. We have given little or no thought to systematically changing these factors to achieve our ends. Our behavior is similar to that of an auto mechanic who says, after a $300 repair bill, "Yes, I know it still

misses, but I followed the instructions in the manual!" Educational practices are compromises with many circumstances, not the least of which is their "wearability" to teachers and administrators. Colleges, like most institutions, settle into routines which produce the fewest problems for those who have the most direct and continuous influence in them.

Accreditation: Practices and Problems

Since a large proportion of people in the higher education industry are intellectuals, academic practice is almost always buttressed by exquisite and involved justifications. Many eloquent advocates and an extensive literature reassure the king that he is completely clothed. Only the most persistent reactions from outside the academic community are likely to break through the thicket of academic lore supporting current practice. All too often, accrediting associations elevate these inbred rationalizations into accreditation policies and criteria.

Accreditation's most serious problem seems to result from the following circumstances:

(1) The academy, as an institution, needs systematic feedback on its forms and practices in terms of their consequences.

(2) Accreditation is an extension of the academy; and the academy, if allowed, will create a lore which justifies those practices that accommodate its influential members, without operational regard for the consequences.

For a number of reasons, the traditional lore of academic quality is once again being questioned,[7] with a focus on accreditation policies. We are beginning to understand that educational "prestige" and "effectiveness" are not necessarily the same thing. In this time of abundant, detailed information, widely and rapidly distributed, it is increasingly difficult to maintain an "image" which diverges from "reality." Furthermore, it is becoming common knowledge that quality outcomes (graduates) are more due to quality inputs (admitted freshmen) than to means of treatment (instruction). I know young horse trainers who have more than adequate stables, equipment, technical skills, and willingness to work, but who still do not consistently produce top horses. Why? They do not get the "Harvard and Yale colts." Similarly, the accreditation process cannot base its assessment of institutions on either image or national rankings of graduates.

At the same time, accreditation cannot rest on traditionally acceptable practices because several cats are already out of the bag.

Cat I—Wang's "unbundling" is on its way.[8] The informed public is aware that examining, instruction, guidance, and structure are separable functions.

Cat II—Nontraditional higher education, despite a difficult and uncertain birthing, exists—and a number of people are taking advantage of it. Creditable degrees are being given by institutions through assessment of proficiencies gained without formal instruction and without benefit of full-time faculty.

Cat III—Competency-based education has made many people aware that the "ends" and "means" of instruction are separable.

At the end of the 1800's and the beginning of the 1900's, an institution's accreditation rested to a great extent on the preparation of the students it admitted. An institution could not be accredited as a college if it provided preparatory education as a part of its curriculum. Frankly, one of the surest ways to control for quality is through admissions. It is usually more difficult to weed out inadequate students once they are admitted than never to admit them. It is particularly difficult to deal with inadequate students after they are admitted, in this time of stabilized and declining enrollments. Nevertheless, the old practice of enforcing quality control at the point of admissions is no longer feasible for all but the most prestigious institutions. Society's commitment to equalitarianism is now united with institutions' need to sustain enrollments.

To assure some minimal quality in institutions of higher education, accreditation must increasingly emphasize exit standards. These exit standards cannot be as different as the institutions are different if any common meaning for degrees is to be retained. This is not to say a graduate with a BS in electrical engineering from one institution must know just as much and no more than a graduate with the same degree from another institution. Yet they ought to have some knowledge in common at a minimum level.

If this is to be achieved, institutional self-studies will have to present evidence that graduates are as competent, at a minimum, as their degrees would indicate. In turn, visiting committees should be expected to verify such evidence.[9] Institutions may apply different emphases in different degree programs; this is to be encouraged in a diverse educational system for a pluralistic society. Nevertheless, individual institutions should not be allowed to debase the academic currency by significantly altering the purpose, meaning, and level of competence expected of various degrees.

By now the reader may be convinced that I believe none of the traditional structures and processes of the academy have intrinsic value, that their only real worth lies in their ability to effect preestablished ends. To the contrary, I do believe certain academic conventions have great value, but for purposes other than assurance of quality. We often try to justify certain conventions by saying they promote learning, when we know very well that they serve other purposes. For example, significant faculty influence in grading and curriculum does not necessarily assure a higher, more consistent level of academic achievement than administrative domination of these areas. Rather, the value of faculty influence lies in the pluralism of institutional control their participation provides.

Although tenure, or reasonable assurance of employment cannot be demonstrated to have a direct effect on the quality of student achievement, it helps to make academic life attractive. Along with pluralism of institutional governance, tenure or reasonable assurance of employment is a basic structural characteristic which gives the academy some insulation from its sponsors. Out of long experience, we should have learned the academy serves

a free society best when it is not a direct agency of government. As expressed in one of the papers of the COPA study on accreditation of nontraditional education,[10] a major task of accreditation is to reinforce the academy's relative independence.

Therefore, while there is need for greater attention in accreditation to results or outcomes, structure and process cannot be ignored if the integrity of the academic estate is to be preserved. I would agree with Grover Andrews in the Summary Report of the COPA Project when he recommends to COPA that it take the following initiative:

> That the Council on Postsecondary Accreditation encourage its member accrediting associations to revise their procedures to include a process—performance evaluation system that places emphasis on the assessment of educational accomplishments for accreditation purposes.[11]

In conclusion, "How do you know a good college when you stumble over one?"

(1) Its graduates are as competent at or before graduation as one would expect respective degree holders to be.

(2) The institution behaves in ways that uphold the integrity of the general academic estate.[12]

To these ends, accreditation processes should focus more concern on quality of student achievement, in terms of the stated meaning of the various degrees an institution offers. At the same time, a stamp of accreditation should mean an institution's processes, forms, and structures have been judged against established criteria of academic integrity. Both of these accreditation functions call for the best of educational criticism. They cannot, I believe, be completely reduced to accountability systems.

A Reflective Note

Reformations, restorations, and even revolutions are made of such stuff as concern for *ends* over *means*. After several years of advocating a refocus on results, student achievements rather than structures, I am only now becoming aware of how many apple carts such a reemphasis turns over. One can start with the separation of examining from instructing and then move to curricular reform, in terms of the effects of instruction on student achievement. But this leads to seeing the faculty member primarily as a tutor or coach of students rather than just a teller of physics, history, literature, etc. One might discover students do not understand an era of history by studying its economics in one department, its battles and personalities in another, its literature in another, and its art and music in still another. Funding state colleges according to time students spend in them and accrediting them by the traditional lore of academic practice are questionable practices. Yet, are some questions best left unasked? Are they too traumatic to deal with, and might they hurt more than they help?

Some advocates of educational change appear to behave as if rational, internal reform of educational practice were possible. The assumption seems to be, "If we could just become quite clear about what we want to accomplish and then be very pragmatic about ways to accomplish it, improvement would result." But this assumes that most of those involved in higher education are sincere about what they are doing and are willing to sacrifice and work for improvement in higher education.

Another scenario and other assumptions about change in general and change in education in particular may be closer to reality. Institutions could be compared to bacteria growing on a culture; when the culture changes, the bacteria either adapt or die. Likewise, change in a social institution is not inner-directed, rational redirection but adaption to an altered environment. Those of us who are restless about the status quo perhaps should relax a bit. If the environment or society demands a change, colleges and universities will either adapt or wither. Rather than trying to engineer or even insist on change, it may be our responsibility to do only the following:

(1) Be reliable sensing agents. That is, explain and discusss our perceptions of the needed changes to our more insulated colleagues.

(2) Become adept in discriminating among institutions which have dinosaurian characteristics, versus those which are more adaptive.

My impression is that society is tugging on education's reins. Our structures and processes are not yielding all we claim they are yielding. It is time to taste the broth. We as educators can either participate in the tasting and the changing of the recipe, or others will do so, as they already are in elementary and secondary education through statewide testing.

Addendum: Critical Characteristics of an Accreditable Institution*

(1) The institution has explicit, comprehensive, and consistent goals and objectives subject to periodic review and revision. (p. 66)

(2) The institution is relatively independent and has significant lay representation on its governing board. Its executive leadership has experience and education commensurate with the level and purpose of the institution and evidences commitment to the values and heritage of the academy. There is a system of operational checks and balances which is effective in self-correcting the institution in terms of its purposes and is particularly influential in curricular planning and the setting of degree requirements. (p. 74)

(3) The institution requires student achievements for its respective degrees commensurate with the general meaning of those degrees; it has in place a systematic means to assure that students meet the degree requirements; and it provides evidence that its system of quality assurance is valid and faithfully employed. (p. 77)

*These listed characteristics are quoted directly from:
John Harris. "Critical Characteristics of an Accreditable Institution, Basic Purposes of Accreditation, and Nontraditional Forms of Most Concern," *Research Reports of the Project To Develop Evaluative Criteria and Procedures for the Accreditation of Nontraditional Education.* The Council on Postsecondary Accreditation, September 1978. Vol. 2, pp. 62-105.

(4) The institution has adequate human, physical, and fiscal resources to accomplish its purposes as judged by academic peers. (p. 79)

(5) The institution displays evidence of high probability of stability and permanence. (p. 80)

(6) There is a manifest correlation between the institution's program and the needs, interests, and aspirations of its students, and provision is made for those student services which are logically related to the institution's mission as well as the students' educational needs. (p. 81)

(7) The institution manifests its integrity by assessable goals and objectives, full disclosure of its codes of practice, sound fiscal practices, ethical advertising and student recruitment, consistent application of its codes, and an ongoing, systematic means of institutional self-evaluation against the criteria of its stated goals and objectives. (p. 83)

Footnotes

[1]Howard R. Bowen. "Goals, Outcomes, and Academic Evaluation," in *Evaluating Educational Quality: A Conference. Summary,* The Council on Postsecondary Accreditation, 1979, pp. 19-28.

[2]*Ibid.,* pp. 22-23.

[3]*Project To Develop Evaluative Criteria and Procedures for Accreditation of Nontraditional Education.* The Council on Postsecondary Accreditation, 1978.

[4]William Earl Troutt. *The Quality Assurance Function of Regional Accreditation* (unpublished). George Peabody College for Teachers, August 1978.

[5]John Harris. "Baccalaureate Requirements: Attainments or Exposures?" *Educational Record,* Winter, 1972, pp. 59-65.

John Harris and Steve Keller. "Curriculum Innovation: Three Dimensions," *Teaching-Learning Issues,* Learning Research Center, The University of Tennessee, Knoxville, Fall, 1974, *No.* 26.

[6]John Harris. "Program Evaluation Funding: An Alternative to Funding Public Higher Education by Numbers of Students," *Southern Review of Public Administration,* December 1973, pp. 322-331.

John Harris. "A New Day for Assessment in Higher Education," *Educational Record,* Summer, 1978, pp. 268-282.

Grover Andrews and John Harris. "A Perspective on the Accreditation of Nontraditional Higher Education," *Peabody Journal of Education,* Vol. 56, No. 3, April 1978, pp. 167-173.

[7]An even cursory reading of Frederick Rudolph's *History of the Curriculum Since 1664* indicates to me that the curriculum has been insistently questioned over the last three centuries. For example, dropping the requirement of proficiency in Greek for admission was a major upheaval.

[8]William K. S. Wang. "The Unbundling of Higher Education," *Duke Law Journal,* Vol. 73, 1975, pp. 58-72.

[9]John Harris. "A New Day for Assessment in Higher Education," *op cit,* p. 279.

[10]John Harris. "Critical Characteristics of an Accreditable Institution, Basic Purposes of Accreditation, and Nontraditional Forms of Most Concern," *Seven Reports of the Project To Develop Evaluative Criteria and Procedures for the Accreditation of Nontraditional Education.* The Council on Postsecondary Accreditation, 1978.

[11]*Assessing Nontraditional Education* (Summary Report of the *Project To Develop Evaluative Criteria and Procedures for the Accreditation of Nontraditional Education.* The Council on Postsecondary Accreditation, Vol. 1, 1978, p. 128.

[12]The reader is referred to the addendum entitled, "Critical Characteristics of an Accreditable Institution." This addendum is a summary list of a part of John Harris's "Critical Characteristics of an Accreditable Institution, Basic Purposes of Accreditation, and Nontraditional Forms of Most Concern," *op cit.* This list deals more specifically with the identifiable marks of a good college than these two general statements.

Current Issues in Higher Education, 1979

The series of seven monographs may be purchased for $12.50.

(1) **Perspectives on Leadership** $2.00

Rosabeth Moss Kanter, "Changing the Shape of Work: Reform in Academe"

Morton Darrow, "Changing Values: Implications for Major Social Institutions"

Michael Maccoby, "Leadership Needs of the 1980s"

(2) **Faculty Career Development** $2.50

Everett Carll Ladd, Jr., "The Work Experience of American College Professors: Some Data and an Argument"

Roger Baldwin, "Adult and Career Development: What Are the Implications for Faculty?"

B. Claude Mathis, "Academic Careers and Adult Development: A Nexus for Research"

Robert T. Blackburn, "Academic Careers: Patterns and Possibilities"

Barbara Lazarus and Martha Tolpin, "Engaging Junior Faculty in Career Planning: Alternatives to the Exit Interview"

Allan O. Pfnister, Jill Solder, and Nina Verroca, "Growth Contracts: Viable Strategy for Institutional Planning Under Changing Conditions?"

(3) **Premises of Federal Student Assistance** $1.50

David W. Breneman, "Introduction"

Kenneth M. Deitch, "Why Strains on Need-Based Financial Aid Are Increasing and Some Options for Relieving Them"

Martin Kramer, "How Good Is the Student Financial Aid System?"

Marshall S. Smith, "Premises of Student Assistance: It Isn't Time for a New Look"

(4) **Employment Priorities in Academe** $2.00

Robert H. Linnell, "Age, Sex, and Ethnic Trade-Offs in Faculty Employment: You Can't Have Your Cake and Eat It Too"

Bernice Resnick Sandler, "You've Come a Long Way, Maybe—Or Why It Still Hurts To Be a Woman in Labor"

Frank Newman, "Can There Be Anything Affirmative About Affirmative Action in the 1980s?"

(5) **Assessment** $2.50

David G. Winter, "Defining and Measuring the Competencies of a Liberal Arts Education"

Alice J. Irby, "Statewide Assessment of Basic Skills"

Thomas R. Plough, "Identifying and Evaluating Major Elements in the Quality of Student Life"

Edward A. Dougherty, "What Is the Most Effective Way To Handle Program Discontinuance?"

John Harris, "How Do You Know a Good College When You Stumble Over One?"

(6) **Tenure** $2.00

Walter P. Metzger, "The History of Tenure"

James O'Toole, "The Poverty of Historicism: A Critique of Walter Metzger's Defense of Tenure"

Penina M. Glazer, "The Concept of Tenure and an Alternative"

(7) **PhDs in Nonacademic Careers: Are There Good Jobs?** $2.00 (2 papers)

Lewis C. Solmon

Robots or Reinsmen: Job Opportunities and Professional Standing for Collegiate ... trators in the 1980s

Robert A. Scott

Order Form

Name _____

Address _____

_____

_____

____ copies of (1)
____ copies of (2)
____ copies of (3)
____ copies of (4)
____ copies of (5)
____ copies of (6)
____ copies of (7)
____ copies of the entire series

I enclose $_____ (payable to AAHE).

Payment must accompany all orders. Orders of less than four papers must be accompanied by a $1.00 special-handling fee. Please make check payable to AAHE. Send order to: Publications Department, American Association for Higher Education, One Dupont Circle, Suite 780, Washington, DC 20036.